

APPLICANT: Leong, Kahan
APPLICANT: Raitano, Arthur B.
APPLICANT: Saffran, Douglas C.
APPLICANT: Uabovolis, Aya
TITLE OF INVENTION: BPC-1: A SECRETED BRAIN-SPECIFIC PROTEIN EXPRESSED AND
FILE REFERENCE: 1703-017 US1
CURRENT APPLICATION NUMBER: US/09/374,135
CURRENT FILING DATE: 1999-08-10
PRIOR APPLICATION NUMBER: 60/095,982
PRIOR FILING DATE: 1998-08-10
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 3
LENGTH: 115
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-374-135-3

Query Match 30.2%; Score 260; DB 3; Length 115;
Best Local Similarity 43.5%; Pred. No. 1,1e-24;
Matches 50; Conservative 20; Mismatches 37; Indels 8; Gaps 2;

QY 52 IFTSPNYSKYPPDRECIYIIIEAPROCIIELYFDEKYSIEPSW-----ECKFDHIEVR 104
DB 1 IFTSPNFPDRYPPIDCVIHSRPHDYYVVFHVFHISTYDKIDAGECEPDPFIETFR 60

QY 105 DGRFGSPITIGRFGCGQNP-VIKSSGRFLMIKFFAGDELSMGFSARVFTPRK 158
DB 61 DGRGSPILARFGDMPRREIRAVSGFLMIRRSDBMLEYQGSFAVAIVSRK 115

RESULT 3

US-08-866-650-3
Sequence 3, Application US/08866650
Patent No. 5939321
GENERAL INFORMATION:
APPLICANT: Greenspan, Daniel S
APPLICANT: Takahara, Kazuhiko
APPLICANT: Hoffman, Guy G
TITLE OF INVENTION: Mammalian Tolloid-Like Protein
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Charles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/866,650
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Berson, Bennett J
REGISTRATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 960296,93839
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1013 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-866-650-3

Query Match 27.7%; Score 238.5; DB 2; Length 1013;
Best Local Similarity 39.3%; Pred. No. 1.2e-20;
Matches 42; Conservative 25; Mismatches 39; Indels 1; Gaps 1;

QY 47 HAEGIFTSPNYSKYPPDRECIYIIIEAPROCIIELYFDEKYSIEPSWCKPDIIEVRDQ 106
DB 779 HSPSGILTSPPNPDKYSRREKCTWVISAIRGHITLAFNE-FVEYQHQCAVYDHLFIPTDG 837

QY 107 PGRFSPITIGRFGCGQNPVIVKSSGRFLMIKFFAGDELSMGFSARV 153
DB 838 ETEKSPITIGRLCGSKIPDPLMATGNEMFIRFISDASVQRKGFOATHS 884

RESULT 4

US-09-021-287-3
Sequence 3, Application US/09021287
Patent No. 5981717
GENERAL INFORMATION:
APPLICANT: Greenspan, Daniel S
APPLICANT: Takahara, Kazuhiko
APPLICANT: Hoffman, Guy G
TITLE OF INVENTION: Mammalian Tolloid-Like Protein
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Charles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/021,287
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/866,650
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Berson, Bennett J
REGISTRATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 960296,93839
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1013 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-021-287-3

Query Match 27.7%; Score 238.5; DB 2; Length 1013;
Best Local Similarity 39.3%; Pred. No. 1.2e-20;
Matches 42; Conservative 25; Mismatches 39; Indels 1; Gaps 1;

QY 47 HAEGIFTSPNYSKYPPDRECIYIIIEAPROCIIELYFDEKYSIEPSWCKPDIIEVRDQ 106
DB 779 HSPSGILTSPPNPDKYSRREKCTWVISAIRGHITLAFNE-FVEYQHQCAVYDHLFIPTDG 837

QY 107 PGRFSPITIGRFGCGQNPVIVKSSGRFLMIKFFAGDELSMGFSARV 153
DB 838 ETEKSPITIGRLCGSKIPDPLMATGNEMFIRFISDASVQRKGFOATHS 884

RESULT 5
US-09-240-473-3
Sequence 3, Application US/09240473

[illegible]

MOLECULE TYPE: protein
US-09-021-287-5

Query Match 27.1%; Score 233.5; DB 2; Length 1013;
Best Local Similarity 40.2%; Pred. No. 4.9e-20;
Matches 43; Conservative 23; Mismatches 40; Indels 1; Gaps 1;

47 HAEGLFTSPNPKYPPDECIYIIEAARQCIELYFDEKYSIEPSWECKPDHIEVRDG 106
779 HSPSGLTSPNWDKPSRRECTWEISATPGHRIKLAFSF-FEIEQHQCAYDHLVEFDG 837
107 PGFSPITGRFCGQONPPVTKSGRFLWIKFPADGELIESGFSARYN 153
838 ETEKSPILGRICGNKIPDPLVATGNMCFVRFVSDASVQKGFQATHS 884

RESULT 8

US-08-991-408-2
Sequence 2, Application US/08991408
Patent No. 6008017

GENERAL INFORMATION:
APPLICANT: ARLETH, ANTHONY J.
APPLICANT: WILLETTTE, ROBERT N.
APPLICANT: ELSHOURBAGY, NABIL A.
APPLICANT: LI, XIAOTONG
TITLE OF INVENTION: HUMAN CARDIAC/BRAIN TOLLLOID-LIKE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,408
FILING DATE:

CLASSIFICATION:

PRICE APPLICATION DATA:
APPLICATION NUMBER: 60/034,471
FILING DATE: 02-JAN-1997
ATTORNEY/AGENT INFORMATION:

NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: ATG-50038
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1013 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-991-408-2

Query Match 27.1%; Score 233.5; DB 3; Length 1013;
Best Local Similarity 40.2%; Pred. No. 4.9e-20;
Matches 43; Conservative 23; Mismatches 40; Indels 1; Gaps 1;

47 HAEGLFTSPNPKYPPDECIYIIEAARQCIELYFDEKYSIEPSWECKPDHIEVRDG 106
779 HSPSGLTSPNWDKPSRRECTWEISATPGHRIKLAFSF-FEIEQHQCAYDHLVEFDG 837
107 PGFSPITGRFCGQONPPVTKSGRFLWIKFPADGELIESGFSARYN 153

DB 838 ETEKSPILGRICGNKIPDPLVATGNMCFVRFVSDASVQKGFQATHS 884

RESULT 9
US-09-240-473-5
Sequence 5, Application US/09240473
Patent No. 6297011

GENERAL INFORMATION:
APPLICANT: Greenspan, Daniel S
APPLICANT: Takahara, Kazuhiko
APPLICANT: Hoffman, Guy G
TITLE OF INVENTION: Mammalian Tolloid-Like Protein
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/240,473
FILING DATE:

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Berson, Bennett J
REGISTRATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 960296.93839
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1013 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-240-473-5

Query Match 27.1%; Score 233.5; DB 3; Length 1013;
Best Local Similarity 40.2%; Pred. No. 4.9e-20;
Matches 43; Conservative 23; Mismatches 40; Indels 1; Gaps 1;

47 HAEGLFTSPNPKYPPDECIYIIEAARQCIELYFDEKYSIEPSWECKPDHIEVRDG 106
779 HSPSGLTSPNWDKPSRRECTWEISATPGHRIKLAFSF-FEIEQHQCAYDHLVEFDG 837
107 PGFSPITGRFCGQONPPVTKSGRFLWIKFPADGELIESGFSARYN 153
838 ETEKSPILGRICGNKIPDPLVATGNMCFVRFVSDASVQKGFQATHS 884

RESULT 10

US-09-432-473-2
Sequence 2, Application US/09432473
Patent No. 6365715

GENERAL INFORMATION:
APPLICANT: ARLETH, ANTHONY J.
APPLICANT: WILLETTTE, ROBERT N.
APPLICANT: ELSHOURBAGY, NABIL A.
APPLICANT: LI, XIAOTONG
TITLE OF INVENTION: HUMAN CARDIAC/BRAIN TOLLLOID-LIKE PROTEIN
FILE REFERENCE: ATG-50038-D1
CURRENT APPLICATION NUMBER: US/09/432,473
CURRENT FILING DATE: 1999-11-01
EARLIER APPLICATION NUMBER: 08/991,408
EARLIER FILING DATE: 1997-12-16

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 18:15:56 ; Search time 57 Seconds

(without alignments)
783.202 Million cell updates/sec

Title: US-09-887-593a-2

Perfect score: 862
1 MHRSVLHVIYASLIILHLS.....ADGELSMGFSARVNTFGK 158

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*

1: geneseqp1980s:*\n2: geneseqp1990s:*\n3: geneseqp2000s:*\n4: geneseqp2001s:*\n5: geneseqp2002s:*\n6: geneseqp2003as:*\n7: geneseqp2003bs:*\n8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	862	100.0	158	3	AAy82894 Human bra
2	851	98.7	526	6	ABU69144 Human NOV
3	851	98.7	533	5	ABU79587 Human CUB
4	802	93.0	525	7	ADc42850 REMAP pro
5	479	55.6	130	4	ABBI10345 Human CDN
6	479	55.6	130	5	ABP66932 Human pol
7	479	55.6	141	4	AAU19919 Novel hum
8	470	54.5	487	7	ABR11868 Human sec
9	470	54.5	487	7	ADc42852 REMAP pro
10	470	54.5	525	4	AAm93946 Human pol
11	470	54.5	525	4	AAU12228 Human PRO
12	470	54.5	525	4	AAU12228 Human PRO
13	470	54.5	525	4	AAU12228 Human PRO
14	470	54.5	525	6	ABU80926 Human PRO
15	470	54.5	525	6	ABU66626 Human PRO
16	470	54.5	525	6	ABU59707 Novel sec
17	470	54.5	525	6	ABO24897 Human sec
18	470	54.5	525	6	ABU66902 Human sec
19	470	54.5	525	6	ADA45633 Novel hum
20	470	54.5	525	6	ADA26938 Human PRO
21	470	54.5	525	6	ADA76064 Human PRO
22	470	54.5	525	6	ADA18714 Human PRO
23	470	54.5	525	6	ADA61337 Homo sapi
24	470	54.5	525	6	ADBI1922 Novel hum
25	470	54.5	525	6	ADBI27663 Human PRO

26	470	54.5	525	6	ADa86142 Novel hum
27	470	54.5	525	6	ADBI15706 Human PRO
28	470	54.5	525	6	ADA47492 Human PRO
29	470	54.5	525	6	ADA67287 Human PRO
30	470	54.5	525	6	ADB30294 Human PRO
31	470	54.5	525	6	ADa85590 Novel hum
32	470	54.5	525	6	ADA96802 Human PRO
33	470	54.5	525	6	ADA79106 Human PRO
34	470	54.5	525	6	ADa87245 Novel hum
35	470	54.5	525	6	ADBI16447 Human PRO
36	470	54.5	525	6	ADA91539 Novel hum
37	470	54.5	525	6	ADBI14602 Human PRO
38	470	54.5	525	6	ADBI18563 Human PRO
39	470	54.5	525	6	ADA93778 Human PRO
40	470	54.5	525	6	ADBI19674 Novel hum
41	470	54.5	525	6	ADBI12986 Human PRO
42	470	54.5	525	6	ABO43055 Novel hum
43	470	54.5	525	6	ADBI14240 Human PRO
44	470	54.5	525	6	ADBI24473 Human PRO
45	470	54.5	525	6	ADa81997 Human PRO

ALIGNMENTS

RESULT 1
ID AAY82894 standard; protein; 158 AA.
AC AAY82894;
DT 25-JUL-2000 (first entry)
DE Human brain specific protein BPC-1.
XX BPC-1; oncogene; oncogenic; cancer; prostate; bladder; antibody;
XX antisense; vaccine; detection; prognosis; drug screening; human.
XX Homo sapiens.
XX WO200009691-A2.
XX 24-FEB-2000.
XX 10-AUG-1999; 99WO-US018250.
XX PF 10-AUG-1999; 98US-0095982P.
XX PR 10-AUG-1998; 98US-0095982P.
XX (UROC-) UROGENESYS INC.
XX (AFAR/) AFAR D E.
XX (HUBE/) HUBERT R S.
XX (LEON/) LEONG K.
XX (RAIT/) RAITANO A B.
XX (SAFR/) SAFRAN D C.
XX (JAKO/) JAKOBOVITS A.
XX Afar DE, Hubert RS, Leong K, Raitano AB, Saffran DC;
XX Jakobovits A;
XX WPI: 2000-206006/18.
XX N-PSDB; AA293040.
XX New isolated BPC-1 polypeptides, useful for developing products for the
XX diagnosis, staging, prognosis and treatment of cancers, particularly
XX prostate or bladder cancer.
XX Claim 1; Fig 1; 79pp; English.
XX BPC-1 polypeptides and polynucleotides can be used for the detection of
XX BPC-1 polypeptides and polynucleotides in biological samples, this is
XX particularly useful for detecting cancers expressing BPC-1, e.g. prostate
XX cancer or bladder cancer. Antibodies directed against BPC-1 or antisense
XX polynucleotides can be used for treating such cancers. The BPC-1

polypeptides can also be used in vaccines for treating or inhibiting the development of a cancer expressing BpC-1. The polypeptides and polynucleotides can also be used for detection, prognosis, drug screening and predicting susceptibility to developing cancer. The BpC-1 polypeptide comprises a C1B domain which is expressed in prostate and bladder carcinoma cells and which shows sequence similarity with C1B domains from other known proteins. In normal human tissues BpC-1 is only expressed in certain tissues of the brain, however, it is expressed at high levels in prostate cancer cells and bladder cancer cells.

50 Sequence 158 AA;

Query Match	100.0%	Score 862	DB 3	length 158
Best Local Similarity	100.0%	Pred. No.	1	8e-90
Matches 158	Conservative	0	Mismatches	0
			Indels	0
			Gaps	0

QY 1 MHGRSVLHVSLILHLHGATKKGTEKQTSEFQKSVGGTMTKAEGGIFTSPPVPS 60

Db 1 MHGRSVLHVSLILHLHGATKKGTEKQTSEFQKSVGGTMTKAEGGIFTSPPVPS 60

QY 61 KYPPDECTYIIAARQCIELYPDEKYSIEPSMECKFDHLEVDGPPGSPITIGRCGQ 120

Db 61 KYPPDECTYIIAARQCIELYPDEKYSIEPSMECKFDHLEVDGPPGSPITIGRCGQ 120

QY 121 QNPVATKSSGRFLMIKFPADGELSEMGFSARYNFPPGK 158

Db 121 QNPVATKSSGRFLMIKFPADGELSEMGFSARYNFPPGK 158

RESULT 2

ID ABU69144 standard; protein; 526 AA.

DT 02-JUN-2003 (first entry)

Human NOVX polypeptide #19.

KM Human.NOVX; metabolic disorder; diabetes; infectious disease; obesity;
KM anorexia; cancer; cardiovascular disorder; asthma; neurogenesis;
KM neurodegenerative disorder; epilepsy; immune disorder; osteoarthritis;
KM haematopoietic disorder; inflammatory skin disorder; dyslipidemia;
KM haematopoiesis; wound healing; angiogenesis; bacterial infection;
KM viral infection; fungal infection; helminthic infection; atherosclerosis
KM protozoal infection; hypertension.

OS Homo sapiens.

PN WO200290504-A2

PD 14-NOV-2002

02-MAY-2002; 2002WO-US014342.

PR	03-MAY-2001	2001US-02838355
PR	04-MAY-2001	2001US-02868900
PR	07-MAY-2001	2001US-02906087
PR	14-MAY-2001	2001US-02907537
PR	15-MAY-2001	2001US-02911899
PR	16-MAY-2001	2001US-02912439
PR	18-MAY-2001	2001US-02920010
PR	21-MAY-2001	2001US-02923747
PR	22-MAY-2001	2001US-02925878
PR	23-MAY-2001	2001US-02931107
PR	29-MAY-2001	2001US-02941109
PR	30-MAY-2001	2001US-02944347
PR	31-MAY-2001	2001US-02948279
PR	16-JUN-2001	2001US-02968968
PR	17-JUL-2001	2001US-03089091
PR	17-AUG-2001	2001US-03133889
PR	21-AUG-2001	2001US-03139515
PR	21-AUG-2001	2001US-03139876

PR 17-SEP-2001; 2001US-0322701P
PR 17-SEP-2001; 2001US-0322802P
PR 17-SEP-2001; 2001US-0324757P
PR 25-SEP-2001; 2001US-0325314P
PR 27-SEP-2001; 2001US-0325682P
PR 27-SEP-2001; 2001US-0325682P
PR 21-NOV-2001; 2001US-0332129P
PR 03-DEC-2001; 2001US-0336882P
PR 14-DEC-2001; 2001US-0340305P
PR 01-MAY-2002; 2002US-00135858

PA (CURA-) CURAGEN CORP

PI Alsbrook JP, Anderson DM, Boldo FL, Burgess CE, Casman SJ,
PI Chapoval A, Edinger S, Gerlach V, Gorman L, Gunther E, Guo X,
PI Kexuda R, Lepley DM, Li L, Liu X, Malyankar UM, Miller CE,
PI Millet I, Padigar M, Patrujan M, Pena CE, Rieger DK, Shenoy SG,
PI Shinkets RA, Spytek KA, Taupier RJ, Vernet CAM, Voss EZ,
PI Zernusen BD;

DR WPI; 2003-103512/09
DR N-PSDB; ACA10129.

PT New isolated NOXV polypeptides and polymucleotides, useful for preventing, diagnosing or treating NOXV-associated disorders, e.g. PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease, PT asthma, or infections.

PS Claim 2; Page 123; 340pp; English

The invention relates to human NOXV polypeptides and the polynucleotides encoding them. The polypeptides, polynucleotides and antibodies that bind immunospecifically to the polypeptides are useful in the manufacture of a medicament for treating a syndrome associated with a human disease, preferably a NOXV-associated disorder. The sequences are useful for treating, preventing or diagnosing diseases such as metabolic disorders, diabetes, obesity, infectious diseases (viral, bacterial, fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular disorders (e.g. hypertension, atherosclerosis), neurodegenerative disorders (e.g. Alzheimer's disease, Parkinson's disease), epilepsy, immune disorders, osteoarthritis, haematopoietic disorders, inflammatory skin disorders, asthma and various dyslipidaemias. The nucleic acids and polypeptides may also be used as targets for the identification of small molecules that modulate or inhibit e.g. neurogenesis, cell differentiation, cell proliferation, haematopoiesis, wound healing and angiogenesis, and in the generation of antibodies that bind immunospecifically to NOXV substances for use in therapeutic or diagnostic methods. The nucleic acids are further used as hybridisation probes, and in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. Sequences AB069126-AB069171 represent human NOXV polypeptides of the invention

SQ Sequence 526 AA.

Query Match	98.7%	Score	851	DB	6	Length	526
Best Local Similarity	100.0%	Pred. No.	1	7e-88			
Matches	156	Conservative	0	Mismatches	0	Indels	0
						Gaps	0

QY	1	MHGSGVLAHVASLLIIHLSGATKTKGTEKOTTSSETKSVQGGIMTWKHAHGGIIFTS	PNPDS	60
Db	1	MHGSGVLAHVASLLIIHLSGATKTKGTEKOTTSSETKSVQGGIMTWKHAHGGIIFTS	PNPDS	60
QY	61	KYPPEKCIYIIIAAPROCIETLPDEKTSIEPSWCKEFDHIEVDGPFQSPIIERFCGQ		120
Db	61	KYPPEKCIYIIIAAPROCIETLPDEKTSIEPSWCKEFDHIEVDGPFQSPIIERFCGQ		120
QY	121	QNPPIKSSGRFLMIKFPADGELESMGSAARYNFTP		156
Db	121	QNPPIKSSGRFLMIKFPADGELESMGSAARYNFTP		156

RESULT 3

ABB79587
ID ABB79587 standard; protein; 533 AA

XX

AC	AB079587;
XX	
DT	21-OCT-2002 (first entry)
XX	
DE	Human CUB domain-containing protein 39362.
XX	
KW	CUB domain protein, 39362; human; antiarteriosclerotic; cardiant;
KW	cerdiprotective; hypotensive; antirheumatic; antiaerthetic;
KW	antiallergic; neuroprotective; cytostatic; antithyroid; vasotrophic;
KW	antialstematic; antidiabetic; nephrotropic; antiparkinsonian;
KW	antiinflammatory; gene therapy.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Domain
FT	/note="predicted N-terminal extracellular domain"
FT	1..344
FT	/label= Signal_peptide
FT	23..25
FT	/note="predicted protein kinase C phosphorylation site"
FT	24..533
FT	/label= Mature_protein
FT	24..27
FT	/note="predicted CAMP- and cGMP-dependent protein kinase phosphorylation site"
FT	27..29
FT	/note="predicted protein kinase C phosphorylation site"
FT	31..34
FT	/note="predicted casein kinase II phosphorylation site"
FT	35..37
FT	/note="predicted protein kinase C phosphorylation site"
FT	41..152
FT	/note="predicted CUB domain"
FT	50..55
FT	/note="predicted N-myristoylation site"
FT	129..131
FT	/note="predicted protein kinase C phosphorylation site"
FT	149..151
FT	/note="predicted protein kinase C phosphorylation site"
FT	172..284
FT	/note="predicted CUB domain"
FT	177..182
FT	/note="predicted N-myristoylation site"
FT	195..198
FT	/note="predicted casein kinase II phosphorylation site"
FT	241..244
FT	/note="predicted casein kinase II phosphorylation site"
FT	274..279
FT	/note="predicted N-myristoylation site"
FT	286..289
FT	/note="predicted casein kinase II phosphorylation site"
FT	290..328
FT	/note="predicted low density lipoprotein receptor class A domain"
FT	306..309
FT	/note="predicted N-glycosylation site"
FT	313..318
FT	/note="predicted N-myristoylation site"
FT	329..332
FT	/note="predicted CAMP- and cGMP-dependent protein kinase phosphorylation site"
FT	333..336
FT	/note="predicted casein kinase II phosphorylation site"
FT	340..343
FT	/note="predicted N-glycosylation site"
FT	341..351
FT	/note="predicted prokaryotic membrane lipoprotein lipid attachment site"
FT	343..348
FT	/note="predicted N-myristoylation site"
FT	345..363
FT	/note="predicted transmembrane domain"
FT	

FT	Domain	364. .533
FT	/note= "predicted C-terminal cytoplasmic domain"	
FT	Modified-site	377. .380
FT	/note= "predicted casein kinase II phosphorylation site"	
FT	Modified-site	397. .399
FT	/note= "predicted protein kinase C phosphorylation site"	
FT	Modified-site	400. .405
FT	/note= "predicted N-myristoylation site"	
FT	Modified-site	421. .424
FT	/note= "predicted cAMP- and cGMP-dependent protein kinases phosphorylation site"	
FT	Modified-site	424. .426
FT	/note= "predicted protein kinase C phosphorylation site"	
FT	Modified-site	434. .439
FT	/note= "predicted N-myristoylation site"	
FT	Modified-site	439. .441
FT	/note= "predicted protein kinase C phosphorylation site"	
FT	Modified-site	442. .447
FT	/note= "predicted N-myristoylation site"	
FT	Modified-site	446. .449
FT	/note= "predicted N-glycosylation site"	
FT	Modified-site	448. .451
FT	/note= "predicted casein kinase II phosphorylation site"	
FT	Modified-site	448. .450
FT	/note= "predicted protein kinase C phosphorylation site"	
FT	Modified-site	481. .484
FT	/note= "predicted N-glycosylation site"	
FT	Modified-site	502. .504
FT	/note= "predicted protein kinase C phosphorylation site"	
FT	Modified-site	506. .509
FT	/note= "predicted casein kinase II phosphorylation site"	
FT	Modified-site	522. .525
FT	/note= "predicted casein kinase II phosphorylation site"	
FT	Modified-site	529. .532
FT	/note= "predicted N-glycosylation site"	
FT	Modified-site	530. .532
FT	/note= "predicted protein kinase C phosphorylation site"	
FT	Region	531. .533
FT	/note= "predicted microbody C-targeting signal"	
FN	WO200259275-A2.	
PD	01-AUG-2002.	
XX		
XX	08-JAN-2002; 2002WO-US000275.	
PE		
XX		
XX	08-JAN-2001; 2001US-0260286P.	
PR		
XX	(MILL-) MILLENNIUM PHARM INC.	
PA		
XX		
XX	Bandaru R;	
PI		
XX		
XX	WPI; 2002-590825/63.	
DR	N-PSDB; AENB4415.	
XX		
XX	New 39362 polypeptide and nucleic acid molecule, useful for detecting,	
PT	preventing or treating 39362-mediated or -related diseases, e.g.	
PT	atherosclerosis, cancer, and in screening assays, in predictive medicine	
PT	or pharmacogenetics.	
XX		
XX		
PS	Claim 5; Page 119; 131pp; English.	
XX		
CC		
XX	The present sequence is the protein sequence of 39362, a novel human CUB	
CC	domain-containing protein family member. The invention provides isolated	
CC	39362 proteins, fusion proteins, antigenic peptides and anti-39362	
CC	antibodies. It also provides isolated 39362 nucleic acid molecules,	
CC	antisense nucleic acid molecules, recombinant expression vectors, host	
CC	cells, and non-human transgenic animals, a method for identifying a	
CC	compound that binds 39362, a method for modulating 39362 polypeptide	
CC	activity, a method of inhibiting aberrant activity of a 39362-expressing	
CC	cell using a peptide, phosphopeptide, small organic molecule or molecule,	
CC	and a method of treating or preventing a disorder characterised by	
CC	aberrant activity of a 39362-expressing cell, especially a cardiovascular	

cell, by administering a compound that modulates the activity or expression of 39362 nucleic acid, 39362 proteins are useful in treating 39362-mediated or -related diseases, such as conditions involving cardiovascular disorders (e.g., atherosclerosis, myocardial infarction, aneurysm, stroke, hypertension), inflammatory disorders (e.g., rheumatoid arthritis), allergy, multiple sclerosis), and cellular proliferation or differentiation (e.g., cancer, Grave's disease, ischaemic disease). They may also be used in diagnosing, preventing or treating asthma, emphysema, chronic pulmonary disease, nephropathy, diabetes, hyperlipidaemia, infections (e.g. bacterial, viral, parasitic), hepatic injury, Parkinson's disease, or Alzheimer's disease, anxiety or cardiovascular conditions associated with interventional procedures, such as restenosis following angioplasty. 39362 proteins, protein homologues, antibodies and nucleic acids are also useful as screening assays, predictive medicine (e.g. diagnostic assays, prognostic assays, forensic biology, monitoring clinical trials, and pharmacogenetics), and methods of treatment (e.g. therapeutic and prophylactic).

SQ Sequence 533 AA;

Query Match	98.7%	Score 851	DB 5	Length 533
Best Local Similarity	100.0%	Pred. No. 1.7e-88		
Matches 156	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Qy 1 MIHGKRVLHIVASLIILHLSGATKKGTETQTSSTQKSVCCGTWTKHAEGGIFTSPPNYP 60

Db 1 MIHGKRVLHIVASLIILHLSGATKKGTETQTSSTQKSVCCGTWTKHAEGGIFTSPPNYP 60

QY 61 KYPDRECTIYLIAAPRQCIELYFDEKXSIEDSWECKFDHIEVRDGGFGFSPIGRFCGQ 120

Db 61 KYPPRECTIYLIAAPRQCIELYFDEKXSIEDSWECKFDHIEVRDGGFGFSPIGRFCGQ 120

	QY	121 QNPEVIKSSGRFLMIKFADGELESMGFSARYNTP	156
	Dd	121 QNPVTKSGRFLMVKFADFELSMGFSAARYNTP	156

RESULT 4
ADCC42850
ID ADCC42850 standard; protein; 525 AA

AC	ADDC42850;
XX	
DT	18-DEC-2003 (first entry)

DE REMAP protein #10.

KW Cytostatic; Antiarteriosclerotic; Anti-HIV; Antiinflammatory;
KW Antiallergic; Antidiabetic; REMAP; pathogenesis.

OS Homo sapiens.

PN WO2003027228-A2.

PD 03-APR-2003

PF 16-JUL-2002; 2002WO-US022833.

PR 17-JUL-2001; 2001US-0306020E
PR 27-JUL-2001; 2001US-0308172P
PR 02-AUG-2001; 2001US-0309702D
PR 10-AUG-2001; 2001US-0311476P
PR 10-AUG-2001; 2001US-0311551P
PR 10-AUG-2001; 2001US-0311718P
PR 24-AUG-2001; 2001US-0314788P
PR 31-AUG-2001; 2001US-0316639P
PR 07-SEP-2001; 2001US-0317996P

PA (INCY-) INCYTE GENOMICS INC.

PI Lal PG, Honchell CD, Forsythe JF, Walia NK, Tang TY, Borowsky ML;
PI Barroso I, Yue H, Warren BA, Thangavelu K, Gietzen KU, Azimzai Y;
PI Lee EA, Baughn MR, Gorvad AE, Duggan BM, Tran B, Li JX;

PI Richardson TW, Elliott VS, Zebrafadian Y, Tran UK, Yao MG,
PI Peterson DP, Luo W, Lehr-Mason PM;
XX
DR WPI; 2003-421156/39.

New human receptors and membrane-associated proteins (REMP), useful for diagnosing, treating or preventing disorders associated with aberrant PT REMAP expression, e.g. cancer, AIDS, atherosclerosis, hypertension or stroke.

PS Claim 1; SEQ ID NO 10; 115pp; English.

The present invention relates to an isolated polypeptide. The polypeptides and polynucleotides are useful in diagnosing, treating and preventing disorders associated with aberrant expression of REMAP, such as cell proliferative, autoimmune/inflammatory, renal, neurological, cardiovascular, metabolic, developmental, endocrine, muscle, gastrointestinal, lipid metabolism or transport disorders, and viral infections. These are also useful in assessing the effects of exogenous compounds on the expression of nucleic acids and amino acid sequences o REMAP, in facilitating drug discovery process, and in investigating the pathogenesis of diseases or medical conditions. Expression and purification were achieved using bacterial or virus-based expression systems. The present sequence represents an REMAP protein of the invention.

SQ Sequence 525 AA;

Query Match	93.0%	Score 802; DB 7;	length 525;
Best Local Similarity	99.3%	Pred. No. 7e-83;	
Matches 146; Conservative	1;	Mismatches 0;	Gaps 0;

QY 10 IVASIIILHLHSGATKKTGKOTTSFQKSVCCGTTWTHGAEGGI FTSINYPSPKSPDPDECI 69
:
2 VVASIIILHLHSGATKKTGKOTTSFQKSVCCGTTWTHGAEGGI FTSINYPSPKSPDPDECI 61

QY YIIKAPRCICELYPDEKYSLEPSSWCKFDIIEVRGPGFGSP IIGFCCGQGNPVIKSS 129

70 YIIKAPRCICELYPDEKYSLEPSSWCKFDIIEVRGPGFGSP IIGFCCGQGNPVIKSS 129

Db 62 YIIKAPRCICELYPDEKYSLEPSSWCKFDHIEVRGPGFGSP IIGFCCGQGNPVIKSS 129

Qy 130 GRFLWIKFFADGELSMGFSARYNFTP 156
122 GRFLWIKFFADGELSMGFSARYNFTP 148
Db

```

RESULT 5
ABBI0345
ID  ABBI0345 standard; protein; 130 AA

```

AC ABB10345;

DT 10-JAN-2002 (first entry)

DE Human CDNA SEQ ID NO: 653

KM Human; gene therapy; neural disorder; immune system disorder;
KM muscular disorder; reproductive disorder; gastrointestinal disorder
KM pulmonary disorder; cardiovascular disorder; renal disorder;
KM proliferative disorder; inflammation.

OS Homo sapiens

PN WO200154474-A2

PD 02-AUG-2001

PF 17-JAN-2001; 2001WO-US001349.

PR-31-JAN-2000; 2000US-0179065P
PR-04-FEB-2000; 2000US-0180628P
PR-24-FEB-2000; 2000US-0184664P
PR-02-MAR-2000; 2000US-0186350P
PR-16-MAR-2000; 2000US-0189874P

CC DNAs. These can be used in the treatment of neural, immune system,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal and proliferative disorders and inflammation. The present sequence
CC is a protein of the invention

Sequence 130 AA;

Query Match	55.6%	Score 479;	DB 4;	length 130;
Best Local Similarity	66.1%	Pred. No. 1.2e-46;		
Matches	80;	Conservative	21;	Mismatches 20;
			Indels	0;
			Gaps	0;

```
QY      38 SVOCGWMTKHAEBGIIETSPNYSKPDECECIYLIEAAPROCIETYFPEKXISPSMECK 97
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      9 ATCCGGIWKRSNGHAFASPNPIYDUSYPENNECCIYILIEAAPROIETLTPEHHYIISPSECR 68

QY      98 FDHIEVRDGEFGFSPIIGRFCCOQNPVYIKSSGRELIWIKFPADGELBMSGFSARNYFTPG 157
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      69 FDIHEVRDGEFGFSPLIDRYCGVKSPPLIRSTGRFMWIKFSSDBEIBGLGFRAKYSFIPTG 120

QY      158 K 158
Db      129 K 129
```

RESULT 6
ABP66932
ID ABP66932 standard; protein; 130 AA.
...

AC	ABP66932;
XX	
DT	09-DEC-2002 (first entry)
DT	

Human polypeptide SEQ ID NO 653.

KM immunotrophic; neuroprotection; cytosolic; dermatological; virucide;
 KM immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnereary;
 KM antiparkinsonian; antischlicking; antiamebic; antifibrillic; cancer;
 KM antipneumatic; hepatotropic; cerebroprotective; antiinflammatory;
 KM antiallergic; antidiabetic; antiliver; anticonvulsant; antifungal;
 KM antiparastatic; cardiac; immune disorder; cardiovascular disorder;
 KM neurological disease; infection; nephrotropic; gene therapy; vaccine.

OS Homo sapiens.

PN US2002090672-A1

PD 11-JUL-2002

PF 17-JAN-2001; 2001US-00764853

PR	31-JAN-2000	2000US-0179665P
PR	04-FEB-2000	2000US-0180528P
PR	28-JUN-2000	2000US-0214868P
PR	07-JUL-2000	2000US-0216647P
PR	07-JUL-2000	2000US-0216680P
PR	11-JUL-2000	2000US-0217487P
PR	11-JUL-2000	2000US-0217966P
PR	14-JUL-2000	2000US-0218290P
PR	26-JUL-2000	2000US-0220963P
PR	26-JUL-2000	2000US-0220964P
PR	14-AUG-2000	2000US-0224519P
PR	14-AUG-2000	2000US-0224519P
PR	14-AUG-2000	2000US-0225676P
PR	14-AUG-2000	2000US-0225688P
PR	14-AUG-2000	2000US-0225720P
PR	14-AUG-2000	2000US-0225747P
PR	14-AUG-2000	2000US-0225758P
PR	14-AUG-2000	2000US-0225758P
PR	30-AUG-2000	2000US-0228982P
PR	30-AUG-2000	2000US-0228982P
PR	01-SEP-2000	2000US-0229287P
PR	01-SEP-2000	2000US-0229333P
PR	01-SEP-2000	2000US-0229344P

PR	01-SEP-2000	2000US-0229345P
PR	05-SEP-2000	2000US-0229509P
PR	05-SEP-2000	2000US-0229513P
PR	08-SEP-2000	2000US-0231413P
PR	21-SEP-2000	2000US-0234423P
PR	21-SEP-2000	2000US-0234474P
PR	25-SEP-2000	2000US-0234497P
PR	27-SEP-2000	2000US-0235634P
PR	29-SEP-2000	2000US-0236327P
PR	29-SEP-2000	2000US-0236367P
PR	29-SEP-2000	2000US-0236368P
PR	29-SEP-2000	2000US-0236396P
PR	29-SEP-2000	2000US-0236370P
PR	02-OCT-2000	2000US-0236682P
PR	02-OCT-2000	2000US-0237033P
PR	02-OCT-2000	2000US-0237038P
PR	02-OCT-2000	2000US-0237039P
PR	02-OCT-2000	2000US-0237040P
PR	13-OCT-2000	2000US-0239935P
PR	20-OCT-2000	2000US-0240960P
PR	20-OCT-2000	2000US-0241785P
PR	20-OCT-2000	2000US-0241809P
PR	01-NOV-2000	2000US-0244611P
PR	17-NOV-2000	2000US-0249297P
PR	08-DEC-2000	2000US-0251856P
PR	08-DEC-2000	2000US-0251868P
PR	08-DEC-2000	2000US-0251869P

PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.

PI Rosen CA, Ruben SM, Barash SC;

DR WPI; 2002-681727/73.
DR N-PSDB; ABV83904.

Novel polypeptide useful for diagnosis, prognosis, prevention, and treatment of immune, hyperproliferative, renal, respiratory, cardiovascular, reproductive, endocrine, gastrointestinal and neurological disorders.

PS Claim 11; SEQ ID NO 653; 369pp + Sequence Listing; English.

The invention relates to novel genes (ABY833682-ABV84101) and proteins (ABP667129) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune hemolytic anaemia, autoimmune chryoiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, Rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WHO at http://www.who.int/pub/published_pat_sequences

SQ Sequence 130 AA;

Query Match	55.6%;	Score 479;	DB 5;	Length 130;
Best Local Similarity	66.1%;	Pred. No. 1.2e-46;		
Matches 80;	Conservative 21;	Mismatches 20;	Indels 0;	Gaps 0;

```
QY 38 SVQCGTWKHAAGGIFTSNVPSKPPDEDCIYIIAARQCELEFDEKNSLEPSWECK 97
      : | | | : : | | | | | : | | | | | | | | | | | | | | | |
Db 9 AATCGGWTARTSGHHASPNYDPSYNNKECIYIIAARQRRLTFLFDEHYITEPSFECR 68
QY 98 FDIHIEVRDGPFGFSPILIRFCGQONPPVTKSSGRPLMIKFFADGELESMSGFSARYNFTPG 157
```

```
Db      69 FDHLEVRDGPFGFSLIRYGVKSPPLIRSTGRFMWIKFSSDELEGLGFPAKISFTPG 128
Qy      158 K 158
Db      129 K 129

RESULT 7
AAU19919
ID      AAU19919 standard, protein, 141 AA.
XX
AC      AAU19919;
XX
DT      04-DEC-2001 (first entry)
XX
DE      Novel human calcium-binding protein #28.
XX
KW      Human; calcium-binding protein; calcium flux; neurological disease;
KW      immune dysfunction; digestive disorder; neoplastic disease;
KW      blood disorder; infectious disease; gene therapy; immunosuppressive;
KW      anitarrthritic; cytostatic; vasotropic; antibacterial; nootropic;
KW      virucide.
XX
OS      Homo sapiens.
XX
PN      WO200155304-A2.
XX
PD      02-AUG-2001.
XX
PF      17-JAN-2001; 2001WO-US001302.
XX
PR      31-JAN-2000; 2000US-0179065P.
PR      04-FEB-2000; 2000US-0180628P.
PR      24-FEB-2000; 2000US-0184664P.
PR      02-MAR-2000; 2000US-0186350P.
PR      16-MAR-2000; 2000US-0189874P.
PR      17-MAR-2000; 2000US-0190076P.
PR      18-APR-2000; 2000US-0198123P.
PR      19-MAY-2000; 2000US-0205515P.
PR      07-JUN-2000; 2000US-0209467P.
PR      28-JUN-2000; 2000US-0214866P.
PR      30-JUN-2000; 2000US-0215135P.
PR      07-JUL-2000; 2000US-0216647P.
PR      07-JUL-2000; 2000US-0216880P.
PR      11-JUL-2000; 2000US-0217487P.
PR      11-JUL-2000; 2000US-0217496P.
PR      14-JUL-2000; 2000US-0218290P.
PR      26-JUL-2000; 2000US-0220963P.
PR      26-JUL-2000; 2000US-0220964P.
PR      14-AUG-2000; 2000US-0224518P.
PR      14-AUG-2000; 2000US-0224519P.
PR      14-AUG-2000; 2000US-0225213P.
PR      14-AUG-2000; 2000US-0225214P.
PR      14-AUG-2000; 2000US-0225266P.
PR      14-AUG-2000; 2000US-0225267P.
PR      14-AUG-2000; 2000US-0225268P.
PR      14-AUG-2000; 2000US-0225270P.
PR      14-AUG-2000; 2000US-0225757P.
PR      14-AUG-2000; 2000US-0225758P.
PR      14-AUG-2000; 2000US-0225759P.
PR      18-AUG-2000; 2000US-0226279P.
PR      22-AUG-2000; 2000US-0226681P.
PR      22-AUG-2000; 2000US-0226682P.
PR      23-AUG-2000; 2000US-0227182P.
PR      30-AUG-2000; 2000US-0227009P.
PR      01-SEP-2000; 2000US-0228924P.
PR      01-SEP-2000; 2000US-0229287P.
PR      01-SEP-2000; 2000US-0229343P.
PR      01-SEP-2000; 2000US-0229344P.
PR      01-SEP-2000; 2000US-0229345P.
PR      05-SEP-2000; 2000US-0229509P.

PR      05-SEP-2000; 2000US-0229513P.
PR      06-SEP-2000; 2000US-0230437P.
PR      06-SEP-2000; 2000US-0230438P.
PR      08-SEP-2000; 2000US-0231242P.
PR      08-SEP-2000; 2000US-0231243P.
PR      08-SEP-2000; 2000US-0231244P.
PR      08-SEP-2000; 2000US-0231413P.
PR      08-SEP-2000; 2000US-0231414P.
PR      08-SEP-2000; 2000US-0231415P.
PR      08-SEP-2000; 2000US-0232081P.
PR      12-SEP-2000; 2000US-0231968P.
PR      14-SEP-2000; 2000US-0232397P.
PR      14-SEP-2000; 2000US-0232398P.
PR      14-SEP-2000; 2000US-0232399P.
PR      14-SEP-2000; 2000US-0232400P.
PR      14-SEP-2000; 2000US-0233061P.
PR      14-SEP-2000; 2000US-0233063P.
PR      14-SEP-2000; 2000US-0233064P.
PR      14-SEP-2000; 2000US-0233065P.
PR      21-SEP-2000; 2000US-0234223P.
PR      25-SEP-2000; 2000US-0234274P.
PR      25-SEP-2000; 2000US-0234997P.
PR      26-SEP-2000; 2000US-0235484P.
PR      27-SEP-2000; 2000US-0235834P.
PR      27-SEP-2000; 2000US-0235836P.
PR      29-SEP-2000; 2000US-0236327P.
PR      29-SEP-2000; 2000US-0236367P.
PR      29-SEP-2000; 2000US-0236368P.
PR      29-SEP-2000; 2000US-0236369P.
PR      29-SEP-2000; 2000US-0236370P.
PR      02-OCT-2000; 2000US-0236802P.
PR      02-OCT-2000; 2000US-0237037P.
PR      02-OCT-2000; 2000US-0237038P.
PR      02-OCT-2000; 2000US-0237039P.
PR      13-OCT-2000; 2000US-0237040P.
PR      13-OCT-2000; 2000US-0239355P.
PR      13-OCT-2000; 2000US-0239937P.
PR      20-OCT-2000; 2000US-0240960P.
PR      20-OCT-2000; 2000US-0241221P.
PR      20-OCT-2000; 2000US-0241785P.
PR      20-OCT-2000; 2000US-0241786P.
PR      20-OCT-2000; 2000US-0241787P.
PR      20-OCT-2000; 2000US-0241808P.
PR      20-OCT-2000; 2000US-0241809P.
PR      01-NOV-2000; 2000US-0244617P.
PR      08-NOV-2000; 2000US-0246474P.
PR      08-NOV-2000; 2000US-0246475P.
PR      08-NOV-2000; 2000US-0246476P.
PR      08-NOV-2000; 2000US-0246477P.
PR      08-NOV-2000; 2000US-0246478P.
PR      08-NOV-2000; 2000US-0246523P.
PR      08-NOV-2000; 2000US-0246524P.
PR      08-NOV-2000; 2000US-0246525P.
PR      08-NOV-2000; 2000US-0246526P.
PR      08-NOV-2000; 2000US-0246527P.
PR      08-NOV-2000; 2000US-0246528P.
PR      08-NOV-2000; 2000US-0246532P.
PR      08-NOV-2000; 2000US-0246609P.
PR      08-NOV-2000; 2000US-0246610P.
PR      08-NOV-2000; 2000US-0246611P.
PR      17-NOV-2000; 2000US-0249207P.
PR      17-NOV-2000; 2000US-0249208P.
PR      17-NOV-2000; 2000US-0249209P.
PR      17-NOV-2000; 2000US-0249210P.
PR      17-NOV-2000; 2000US-0249211P.
PR      17-NOV-2000; 2000US-0249212P.
PR      17-NOV-2000; 2000US-0249213P.
PR      17-NOV-2000; 2000US-0249214P.
PR      17-NOV-2000; 2000US-0249215P.
PR      17-NOV-2000; 2000US-0249216P.
```

PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0251989P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI, 2001-465568/50.
 DR N-PSDB; AAS31604.
 XX
 PT Isolated nucleic acid molecule encoding a calcium-binding protein is used
 PT in preventing, treating or ameliorating a medical condition.
 XX
 XX Claim 11: SEQ ID NO 116; 542pp; English.
 XX
 XX The present invention relates to the isolation of novel human calcium-
 CC binding proteins, and cDNA (AAS31577-AAS31654) and genomic sequences
 CC encoding for these proteins. The sequences of the invention are useful in
 CC the diagnosis, prevention and/or prognosis of diseases associated with
 CC aberrant calcium flux. Such disorders include neurological diseases (e.g.
 CC amyloidotic lateral sclerosis, ALS), immune dysfunction (e.g. severe
 CC combined immunodeficiency, SCID), digestive disorders (e.g. irritable
 CC bowel syndrome, IBS), neoplastic disease (e.g. cancer), blood disorders
 CC (e.g. haemophilia), and/or infectious disease (e.g. acquired
 CC immunodeficiency syndrome, AIDS). The novel calcium-binding proteins are
 CC also useful as screening tools to identify antagonists and/or agonists
 CC that may enhance or inhibit activities mediated by calcium-binding
 CC proteins. The polynucleotides of the invention are also useful in gene
 CC therapy. AAU19892-AAU19969 represent the novel human calcium-binding
 CC proteins. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 CC
 XX Sequence 141 AA;
 XX
 XX Query Match 55.6%; Score 479; DB 4; Length 141;
 XX Best Local Similarity 66.1%; Pred. No. 1,4e-46;
 XX Matches 80; Conservative 21; Mismatches 20; Indels 0; Gaps 0;

AB011868
 ID ABB11868 standard; peptide; 352 AA.
 XX
 AC ABB11868;
 XX
 DT 11-JAN-2002 (first entry)
 XX
 XX Human secreted protein homologue, SEQ ID NO:2238.
 DE
 XX Human: cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulatory activity;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiaustmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnereary; anticancer.
 XX
 XX Homo sapiens.
 OS
 XX WO200157188-A2.
 XX
 XX 09-AUG-2001.
 PD
 XX 05-FEB-2001; 2001WO-US003800.
 PF
 XX 03-FEB-2000; 2000US-00496914.
 XX 27-APR-2000; 2000US-00560875.
 XX
 XX (HYSR-) HYSRQ INC.
 PA
 XX Tang YT, Liu C, Drmanac RT;
 XX
 XX WPI, 2001-457740/49.
 DR N-PSDB; ABA09112.
 XX
 XX Human proteins and DNA encoding sequences useful for preventing, treating
 PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
 PT and cancer.
 XX
 XX Claim 20; Page 268; 1963pp; English.
 XX
 XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound

XX Claim 8; SEQ ID NO 4135; 1380bp + Sequence Listing; English.
XX
XX The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5' and 3' ends of the cDNA molecules have
CC been determined. Primers for synthesising the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is a polypeptide encoded by a full length
CC human cDNA of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in CD-ROM
CC format directly from EPO
XX
XX Sequence 525 AA;
SQ
Query Match 54.5%; Score 470; DB 4; Length 525;
Best Local Similarity 54.9%; Pred. No. 8.8e-45;
Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;
QY 6 SVLHI--VASLIILHSGATKKGTSEKQTSSEKQSVQCGTWKHAEGIFTSPTNPSKYP 63
DB 8 SVLKXVLLITLVVVEGIAVAKXTODGONIGIKHIPATCGIWRTSNGGHFASPNYDSDYP 67
QY 64 PPRECIYITEAPROCIETLYFDEKYSIESWECKFPHIEVRDGFSPFIIGRCGQGNP 123
DB 68 PKKECIYITEAPROCIETLYFDEHYIEBSFECRFPHLEVRDGFSPFIIDRYCGVSKP 127
QY 124 PVKSSGFLMTKTFPADGELESWGFSAFNFTP 156
DB 128 PLIRSTGRMWMIKFSDELELGIFRAKISFTP 160
RESULT 11
AAU12228
ID AAU12228 standard; protein; 525 AA.
XX
XX AAU12228;
AC
XX
XX 24-OCT-2001 (first entry)
DT
XX
XX Human PRO4401 polypeptide sequence.
DE
XX
XX Human secretory and transmembrane; PRO; mammalian; cancer; lung; breast;
XX prostate; cervical; tumour necrosis factor-alpha; TNF-alpha; cartilage;
XX ear; proliferation; glucose; free fatty acid; skeletal muscle; adipocyte;
XX A-peptide; factor VIIa; gene therapy.
OS
XX Homo sapiens.
PN
XX WO200140466-A2.
PD
XX 07-JUN-2001.
XX
XX 01-DEC-2000; 2000WO-US032678.
XX
XX 01-DEC-1999; 99WO-US028301.
XX 01-DEC-1999; 99WO-US028634.
XX 02-DEC-1999; 99WO-US028551.
XX 02-DEC-1999; 99WO-US028565.
XX 02-DEC-1999; 99US-0170282P.
XX 09-DEC-1999; 99WO-US030911.
XX 16-DEC-1999; 99WO-US030911.
XX 20-DEC-1999; 99WO-US030999.
XX 30-DEC-1999; 99WO-US031243.
XX 30-DEC-1999; 99WO-US031274.
XX 05-JAN-2000; 2000WO-US000219.
XX 06-JAN-2000; 2000WO-US000277.
XX 11-FEB-2000; 2000WO-US003565.

PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005841.
PR 03-MAR-2000; 2000WO-US0187202P.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007332.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 05-JUN-2000; 2000WO-US020832P.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
XX
XX (GENTH) GENENTECH INC.
XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
XX Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
XX Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WT, Zhang Z;
XX WPI: 2001-408281/43.
XX N-PSDB; AAS21300.
XX
XX Isolated, secretory and transmembrane PRO polypeptide used to detect
XX other PRO polypeptides, link bioactive molecules to cells expressing PRO
XX polypeptides, and detect the presence of mammalian tumors e.g. lung,
XX breast, prostate, cervical.
XX
XX Claim 12; Fig 114; 813bp; English.
XX
XX AAU12172-AAU12446 represent novel human secretory and transmembrane PRO
XX polypeptides. The PRO polypeptides are useful to detect other PRO
XX polypeptides, to link bioactive molecules to cells expressing PRO
XX polypeptides, to modulate biological activities of cells expressing PRO
XX polypeptides, and to detect the presence of mammalian lung, colon,
XX breast, prostate, rectal, cervical or liver tumors by comparing PRO
XX polypeptide expression in a cell sample to that in a control sample. Some
XX of the 275 sequences are also useful to stimulate the release of tumour
XX necrosis factor-alpha (TNF-alpha) from human blood, the proliferation or
XX differentiation of chondrocytes, the proliferation or gene expression in
XX pericyte cells, the release of proteoglycans from cartilage, the
XX proliferation of inner ear utricular supporting cells or of T-
XX lymphocytes, the release of a cytokine from peripheral blood monocytes
XX (PBMCs), or the proliferation of endothelial cells. Some of the PRO
XX polypeptides may modulate glucose or free fatty acid uptake by skeletal
XX muscle cells or by adipocytes; or inhibit binding of A-peptide to factor
XX VIIa. The PRO polypeptides can be used in assays to identify molecules
XX involved in binding interactions. The polynucleotides encoding PRO
XX polypeptides can be used to generate probes, antisense RNA/DNA,
XX transgenic or knock out animals and can be used in gene therapy
XX
XX Sequence 525 AA;
SQ
Query Match 54.5%; Score 470; DB 4; Length 525;
Best Local Similarity 54.9%; Pred. No. 8.8e-45;
Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;
QY 6 SVLHI--VASLIILHSGATKKGTSEKQTSSEKQSVQCGTWKHAEGIFTSPTNPSKYP 63
DB 8 SVLKXVLLITLVVVEGIAVAKXTODGONIGIKHIPATCGIWRTSNGGHFASPNYDSDYP 67
QY 64 PPRECIYITEAPROCIETLYFDEKYSIESWECKFPHIEVRDGFSPFIIGRCGQGNP 123

PR 18-FEB-2000; 2000WO-US004341.
 PR 18-FEB-2000; 2000WO-US004342.
 PR 22-FEB-2000; 2000WO-US004343.
 PR 22-FEB-2000; 2000WO-US004344.
 PR 24-FEB-2000; 2000WO-US004911.
 PR 24-FEB-2000; 2000WO-US005004.
 PR 01-MAR-2000; 2000WO-US005005.
 PR 02-MAR-2000; 2000WO-US005006.
 PR 02-MAR-2000; 2000WO-US005007.
 PR 10-MAR-2000; 2000WO-US005008.
 PR 15-MAR-2000; 2000WO-US005009.
 PR 20-MAR-2000; 2000WO-US005010.
 PR 21-MAR-2000; 2000WO-US005011.
 PR 30-MAR-2000; 2000WO-US005012.
 PR 17-MAY-2000; 2000WO-US005013.
 PR 22-MAY-2000; 2000WO-US005014.
 PR 30-MAY-2000; 2000WO-US005015.
 PR 02-JUN-2000; 2000WO-US005016.
 PR 28-JUL-2000; 2000WO-US005017.
 PR 11-AUG-2000; 2000WO-US005018.
 PR 23-AUG-2000; 2000WO-US005019.
 PR 24-AUG-2000; 2000WO-US005020.
 PR 08-NOV-2000; 2000WO-US005021.
 PR 10-NOV-2000; 2000WO-US005022.
 PR 01-DEC-2000; 2000WO-US005023.
 PR 20-DEC-2000; 2000WO-US005024.
 PR 28-DEC-2000; 2000WO-US005025.
 PR 28-FEB-2001; 2001US-00796498.
 PR 28-FEB-2001; 2001US-00796499.
 PR 01-MAR-2001; 2001US-00806520.
 PR 09-MAR-2001; 2001US-00806521.
 PR 14-MAR-2001; 2001US-00806522.
 PR 22-MAR-2001; 2001US-00806523.
 PR 05-APR-2001; 2001US-00806524.
 PR 10-MAY-2001; 2001US-00854208.
 PR 10-MAY-2001; 2001US-00854209.
 PR 18-MAY-2001; 2001US-00860216.
 PR 25-MAY-2001; 2001US-00860217.
 PR 25-MAY-2001; 2001US-00860218.
 PR 25-MAY-2001; 2001US-00860219.
 PR 01-JUN-2001; 2001US-00872035.
 PR 01-JUN-2001; 2001US-00872036.
 PR 05-JUN-2001; 2001US-00874503.
 PR 14-JUN-2001; 2001US-00882636.
 PR 19-JUN-2001; 2001US-00882637.
 PR 20-JUN-2001; 2001US-00882638.
 PR 21-JUN-2001; 2001US-00882639.
 PR 22-JUN-2001; 2001US-00882640.
 PR 29-JUN-2001; 2001US-00882641.
 PR 09-JUL-2001; 2001US-00882642.
 PR 18-JUL-2001; 2001US-00882643.
 PR 06-AUG-2001; 2001US-00882644.
 PR 09-AUG-2001; 2001US-00882645.
 PR 16-AUG-2001; 2001US-00882646.
 PR 19-DEC-2001; 2001US-00882647.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;
 PI Gerlitsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 XX
 DR WPI; 2003-332040/31.
 DR N-PSDB; ACN03659.
 PT
 PT New secreted and transmembrane PRO nucleic acids, useful for gene
 PT therapy, in chromosome and gene mapping, as chromosome markers, in tissue
 PT typing, and in chromosome identification.
 XX
 PS Claim 12; Fig 114; 660pp; English.
 CC
 CC The present invention relates to the isolation of novel human PRO
 CC polypeptides, and the polynucleotide sequences encoding them. The PRO
 CC polypeptides are secreted and transmembrane proteins. The PRO

CC polypeptides are useful for detecting other PRO polypeptides, for linking
 CC bioactive molecules to cells expressing PRO polypeptides, for modulating
 CC biological activities of cells expressing PRO polypeptides, and for
 CC identifying agonists or antagonists. The PRO polypeptides are useful for
 CC for stimulating the release of tumour necrosis factor (TNF)-alpha from
 CC human blood, for stimulating the proliferation or differentiation of
 CC chondrocytes, and detecting the presence of tumours. The polynucleotide
 CC sequences encoding PRO polypeptides are useful as hybridisation probes,
 CC in chromosome and gene mapping, in the generation of antisense RNA and
 CC DNA, in the preparation of PRO polypeptides, for generating transgenic
 CC animals or knockout animals, for the genetic analysis of individuals with
 CC genetic disorders, and in gene therapy. AB06570-AB066844 represent the
 CC human PRO polypeptides of the invention. Note: The sequence data for this
 CC patent was obtained in electronic format directly from the USPTO web site
 CC at segdata.uspto.gov/pspsidentry.html
 CC
 XX

Sequence 525 AA;

Query Match 54.5%; Score 470; DB 6; Length 525;
 Best Local Similarity 54.9%; Pred. No. 8, 8e-45;
 Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;

Qy 6 SVLHI--VASLIILHSGATKKGTEKQTTSERQKVQCGTWKHAEGIFTSNPNSKYP 63
 Db 8 SVLKVLLITVAVESGIAVAQKTKQDQNGINHIIPATQGIWRTSNGGHFASPNPDSYP 67
 Qy 64 PDRECIYIIIEAARQCIILYFDEKYSIPSMCKPDHLEVRDGPFGSPFIIGRFGQGNP 123
 Db 68 PNKECIYIIIEAARQRIILTFDEHYIIPSPFECRDHLEVRDGPFGSPFIIDRYGVKSP 127
 Qy 124 PVKSGRFLWIKFPADGELESMGFSARYNFTP 156
 Db 128 PLIRSTGRFWIKFSDELEGLGRKXSFIP 160

Search completed: May 18, 2004, 18:22:24
 Job time : 60 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: May 18, 2004, 18:19:46 ; Search time 39 Seconds
(without alignments)
1278.253 Million cell updates/sec

Title: US-09-887-593a-2
Perfect score: 862
Sequence: 1 MHGRSVLHIVASLLIHLHS.....ADGELHSMGFSARYNTPPK 158

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	851	98.7	533	4	Q8TDF5	Q8TDF5 homo sapien
2	843	97.8	533	4	Q86W85	Q86W85 homo sapien
3	818	94.9	533	11	Q8R417	Q8R417 mus musculu
4	818	94.9	533	11	Q8OX39	Q8OX39 mus musculu
5	814	94.4	533	11	Q8C4S3	Q8C4S3 mus musculu
6	802	93.0	156	4	Q8TDF4	Q8TDF4 mus musculu
7	690	80.0	500	4	Q8ND78	Q8ND78 homo sapien
8	494	57.3	466	11	Q8NCM2	Q8NCM2 mus musculu
9	470	54.5	525	4	Q8NC67	Q8NC67 homo sapien
10	468	54.3	491	4	Q7Z381	Q7Z381 homo sapien
11	465	53.9	444	11	Q8C4O8	Q8C4O8 mus musculu
12	465	53.9	525	11	Q8BNC6	Q8BNC6 mus musculu
13	341	39.6	451	4	Q8ND51	Q8ND51 homo sapien
14	260	30.2	321	5	Q61849	Q61849 caenorhabdi
15	250	29.0	677	5	Q9VVC7	Q9VVC7 drosophila
16	238.5	27.7	1013	11	Q62381	Q62381 mus musculu

17	233.5	27.1	1008	13	Q9DER7	Q9DER7 gallus gall
18	233.5	27.1	1013	4	Q9NQ54	Q9NQ54 homo sapien
19	233.5	27.1	1013	4	Q43897	Q43897 homo sapien
20	229.5	26.6	574	5	Q86RL8	Q86RL8 illyanassa o
21	228.5	26.5	1012	11	Q9WVW6	Q9WVW6 mus musculu
22	227.5	26.4	1007	13	Q8J128	Q8J128 xenopus lae
23	227.5	26.4	3623	4	Q60494	Q60494 homo sapien
24	227.5	26.4	3623	11	Q70244	Q70244 ratius norv
25	227	26.3	926	4	Q9UQ00	Q9UQ00 homo sapien
26	227	26.3	977	13	Q91925	Q91925 xenopus lae
27	227	26.3	1015	4	Q9Y6L7	Q9Y6L7 homo sapien
28	226	26.2	735	13	Q57381	Q57381 xenopus lae
29	226	26.2	926	11	Q8QZT7	Q8QZT7 mus musculu
30	226	26.2	1019	13	Q57382	Q57382 xenopus lae
31	225	26.1	145	11	Q8BP20	Q8BP20 mus musculu
32	224.5	26.0	923	13	Q8OPX6	Q8OPX6 brachydanio
33	224.5	26.0	923	13	Q8AXP1	Q8AXP1 brachydanio
34	224	26.0	555	4	Q9H2E2	Q9H2E2 homo sapien
35	224	26.0	901	4	Q9H2E4	Q9H2E4 homo sapien
36	224	26.0	901	4	Q9H2D5	Q9H2D5 homo sapien
37	224	26.0	906	4	Q9H2D4	Q9H2D4 homo sapien
38	224	26.0	906	4	Q9H2E3	Q9H2E3 homo sapien
39	221	25.6	609	4	Q96190	Q96190 homo sapien
40	221	25.6	644	4	Q961H5	Q961H5 homo sapien
41	221	25.6	704	4	Q9H2E1	Q9H2E1 homo sapien
42	221	25.6	905	13	Q8U0L4	Q8U0L4 gallus gall
43	221	25.6	919	13	Q8UVR0	Q8UVR0 gallus gall
44	221	25.6	923	4	Q86T59	Q86T59 homo sapien
45	221	25.6	936	13	Q8UVQ9	Q8UVQ9 gallus gall

ALIGNMENTS

RESULT 1

Q8TDF5

AC Q8TDF5;

DT 01-JUN-2002 (TREMURel. 21, Created)

DT 01-JUN-2002 (TREMURel. 21, Last sequence update)

DT 01-OCT-2003 (TREMURel. 25, Last annotation update)

DE Neuropilin and tolloid like-1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Retina;

RX MEDLINE=21940629; PubMed=11943477;

RA Stohr H., Berger C., Frohlich S., Weber B.H.;

RT "A novel gene encoding a putative transmembrane protein with two extracellular CUB domains and a low-density lipoprotein class A

RT module: isolation of alternatively spliced isoforms in retina and brain.";

RT Gene 286:223-231 (2002).

CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.

DR EMBL; AF448838; AAM18026.1; --

DR Genew; HGNC:13823; NEROL.

DR InterPro; IPR000859; CUB.

DR InterPro; IPR002172; LDL_receptor_A.

DR Pfam; PF00431; CUB; 2.

DR Pfam; PF00057; ldl_recept_a; 1.

DR SMART; SM00042; CUB; 2.

DR SMART; SM00192; LDLa; 1.

DR PROSITE; PS01180; CUB; 2.

DR PROSITE; PS50068; LDLRA_2; 1.

SQ SEQUENCE 533 AA; 60218 MW; 5B3BBE7BB8031B1 CRC64;

Query Match 98.7%; Score 851; DB 4; Length 533;

Best Local Similarity 100.0%; Pred. No. 1,1e-85;

Matches 156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 M1HGRSVLHIVASLIILHLSGATKKGTETKOTTSSTOKSVQCGTWTAKHAGGIFTSNNYPS 60
DB 1 M1HGRSVLHIVASLIILHLSGATKKGTETKOTTSSTOKSVQCGTWTAKHAGGIFTSNNYPS 60
QY 61 KYPPDRECIYIIIEAAROCIEIYFDEKYSIIEPSWECKPFHIEVRDGPFGSPFIIGRFGCG 120
DB 61 KYPPDRECIYIIIEAAROCIEIYFDEKYSIIEPSWECKPFHIEVRDGPFGSPFIIGRFGCG 120
QY 121 QNPPVTKSGRFLWIKFPADGELSMGFSARYNFTF 156
DB 121 QNPPVTKSGRFLWIKFPADGELSMGFSARYNFTF 156

RESULT 2
Q86W85 PRELIMINARY; PRT; 533 AA.
AC Q86W85;
DT 01-JUN-2003 (TREMELREL. 24, Created)
DT 01-JUN-2003 (TREMELREL. 24, Last sequence update)
DT 01-OCT-2003 (TREMELREL. 25, Last annotation update)
DE Neuropilin (NRP) and toll-like 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxId=9606;
RN (1)
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RU Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC050329; AAH50329.1; -.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR002172; LDL_receptor_A.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00057; ldl_recept_a; 1.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00192; LDLa; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS50068; LDLRA_2; 1.
DR SEQUENCE 533 AA; 60174 MW; DC1B7EBC2C3130E7 CRC64;

Query Match
Best Local Similarity 97.8%; Score 843; DB 4; Length 533;
Matches 155; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 M1HGRSVLHIVASLIILHLSGATKKGTETKOTTSSTOKSVQCGTWTAKHAGGIFTSNNYPS 60
DB 1 M1HGRSVLHIVASLIILHLSGATKKGTETKOTTSSTOKSVQCGTWTAKHAGGIFTSNNYPS 60
QY 61 KYPPDRECIYIIIEAAROCIEIYFDEKYSIIEPSWECKPFHIEVRDGPFGSPFIIGRFGCG 120
DB 61 KYPPDRECIYIIIEAAROCIEIYFDEKYSIIEPSWECKPFHIEVRDGPFGSPFIIGRFGCG 120
QY 121 QNPPVTKSGRFLWIKFPADGELSMGFSARYNFTF 156
DB 121 QNPPVTKSGRFLWIKFPADGELSMGFSARYNFTF 156

RESULT 3
Q8R417 PRELIMINARY; PRT; 533 AA.
AC Q8R417;
DT 01-JUN-2002 (TREMELREL. 21, Created)
DT 01-JUN-2002 (TREMELREL. 21, Last sequence update)
DT 01-OCT-2003 (TREMELREL. 25, Last annotation update)
DE Neuropilin and toll-like 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxId=10090;
RN (1)
RN SEQUENCE FROM N.A.

```

```

RC TISSUE=Retina;
RX MEDLINE=21940629; PubMed=11943477;
RA Stohr H., Berger C., Frohlich S., Weber B.H.;
RT "A novel gene encoding a putative transmembrane protein with two
RT extracellular CUB domains and a low-density lipoprotein class A
RT module; Isolation of alternatively spliced isoforms in retina and
RT brain.";
RL Gene 286:223-231(2002).
CC -1 SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AF448840; AAM18028.1; -.
DR MGI; MGI:2180216; Nocol.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR002172; LDL_receptor_A.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00057; ldl_recept_a; 1.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00192; LDLa; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS50068; LDLRA_2; 1.
DR SEQUENCE 533 AA; 60214 MW; 6375B3972F8D558F CRC64;

Query Match
Best Local Similarity 94.9%; Score 818; DB 11; Length 533;
Matches 147; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 M1HGRSVLHIVASLIILHLSGATKKGTETKOTTSSTOKSVQCGTWTAKHAGGIFTSNNYPS 60
DB 1 M1HGRSVLHIVASLIILHLSGATKKGTETKOTTSSTOKSVQCGTWTAKHAGGIFTSNNYPS 60
QY 61 KYPPDRECIYIIIEAAROCIEIYFDEKYSIIEPSWECKPFHIEVRDGPFGSPFIIGRFGCG 120
DB 61 KYPPDRECIYIIIEAAROCIEIYFDEKYSIIEPSWECKPFHIEVRDGPFGSPFIIGRFGCG 120
QY 121 QNPPVTKSGRFLWIKFPADGELSMGFSARYNFTF 156
DB 121 QNPPVTKSGRFLWIKFPADGELSMGFSARYNFTF 156

RESULT 4
Q8OX39 PRELIMINARY; PRT; 533 AA.
AC Q8OX39;
DT 01-JUN-2003 (TREMELREL. 24, Created)
DT 01-JUN-2003 (TREMELREL. 24, Last sequence update)
DT 01-OCT-2003 (TREMELREL. 25, Last annotation update)
DE Neuropilin-and toll-like protein 1 (Brain-specific transmembrane
DE protein BTCL1).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxId=10090;
RN (1)
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feigold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Burow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo W.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Tohilyuk S., Carninci P., Prange C.,
RA Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butcherfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

```


RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX PubMed:12810072;
 RA Michishita M., Ikeda T., Nakashiba T., Ogawa M., Tashiro K., Honjo T.,
 RA Doi K., Itohara S., Endo S.;
 RT "A novel gene, Brc1l, encoding CUB and LDLa domains is expressed in
 RT restricted areas of mouse brain."
 RL Biochem. Biophys. Res. Commun. 306:680-686 (2003).
 DR EMBL: BC051145; AAH51145.1; -
 DR EMBL: AY138990; AAH38318.1; -
 DR InterPro: IPR000859; CUB;
 DR InterPro: IPR002172; LDL_receptor_A.
 DR Pfam: PF00431; CUB; 2.
 DR Pfam: PF00057; ldl_recept_a; 1.
 DR SMART: SM00042; CUB; 2.
 DR SMART: SM00192; LDLa; 1.
 DR PROSITE: PS01180; CUB; 2.
 DR PROSITE: PS50068; LDLRA_2; 1.
 DR Transmembrane.
 KM SEQUENCE 533 AA; 60242 MW; 6375B395421D558F CRC64;
 SQ
 Query Match 94.9%; Score 818; DB 11; Length 533;
 Best Local Similarity 94.2%; Pred. No. 4.9e-82;
 Matches 147; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 QY 1 MHGSRVLIHVASLIILHLSGATKKGTEKQTTSRTQKSVQCGTWTKAEGGIFTSPTNPS 60
 DB 1 MYGSLPHIILASIIILHSGATKKGTEKQTTPETQKSVQCGTWTKAEGGVFTSPNPS 60
 QY 61 KYPDPRECIYIIIEAPROCIIEFYDEKYSIEPSWECKFDHIEVRDGPFGFSPIIGRFGCG 120
 DB 61 KYPDPRECVYIIIEAPROCIIEFYDEKYSIEPSWECKFDHIEVRDGPFGFSPIIGRFGCG 120
 QY 121 QNPPVTKSSGRFLWIKFPADGELSMGFSARYNFTF 156
 DB 121 QNPPVTKSSGRFLWIKFPADGELSMGFSARYNFTF 156
 RESULT 5
 O8C4S3 PRELIMINARY; PRT; 533 AA.
 AC O8C4S3;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hypothetical Microbodies C-terminal targeting signal/LDL-receptor
 DE class A.
 GN NETOL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573 (2002).
 DR EMBL: AK081325; BAC38196.1; -
 DR MGD: MGI:2180216; Netol.
 DR InterPro: IPR000859; CUB.
 DR InterPro: IPR002172; LDL_receptor_A.

DR Pfam: PF00431; CUB; 2.
 DR SMART: SM00057; ldl_recept_a; 1.
 DR SMART: SM00042; CUB; 2.
 DR PROSITE: SM00192; LDLa; 1.
 DR PROSITE: PS01180; CUB; 2.
 DR PROSITE: PS50068; LDLRA_2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 533 AA; 60248 MW; 04319A6F9456CB3D CRC64;
 Query Match 94.4%; Score 814; DB 11; Length 533;
 Best Local Similarity 93.6%; Pred. No. 1.4e-81;
 Matches 146; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 QY 1 MHGSRVLIHVASLIILHLSGATKKGTEKQTTSRTQKSVQCGTWTKAEGGIFTSPTNPS 60
 DB 1 MYGSLPHIILASIIILHSGATKKGTEKQTTPETQKSVQCGTWTKAEGGVFTSPNPS 60
 QY 61 KYPDPRECIYIIIEAPROCIIEFYDEKYSIEPSWECKFDHIEVRDGPFGFSPIIGRFGCG 120
 DB 61 KYPDPRECVYIIIEAPROCIIEFYDEKYSIEPSWECKFDHIEVRDGPFGFSPIIGRFGCG 120
 QY 121 QNPPVTKSSGRFLWIKFPADGELSMGFSARYNFTF 156
 DB 121 QNPPVTKSSGRFLWIKFPADGELSMGFSARYNFTF 156
 RESULT 6
 O8TDF4 PRELIMINARY; PRT; 156 AA.
 AC O8TDF4;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Neutrophilin and collodid like-1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 CC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retina;
 RX MEDLINE=21940629; PubMed=11943477;
 RA Stroth H., Berger C., Frohlich S., Weber B.H.;
 RT extracellular CUB domains and a low-density lipoprotein class A
 RT module: isolation of alternatively spliced isoforms in retina and
 RT brain."
 RL Gene 286:223-231 (2002).
 CC -I- SIMILARITY: CONTAINS 1 CUB DOMAIN.
 DR EMBL: AF448839; AAM18027.1; -
 DR InterPro: IPR000859; CUB.
 DR Pfam: PF00431; CUB; 1.
 DR SMART: SM00042; CUB; 1.
 DR PROSITE: PS01180; CUB; 1.
 SQ SEQUENCE 156 AA; 17669 MW; 25FB8BF4553C2AB CRC64;
 Query Match 93.0%; Score 802; DB 4; Length 156;
 Best Local Similarity 99.3%; Pred. No. 7e-81;
 Matches 146; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 10 IVASLIILHLSGATKKGTEKQTTSRTQKSVQCGTWTKAEGGIFTSPTNPSKYPDPRECI 69
 DB 9 VVASLIILHLSGATKKGTEKQTTSRTQKSVQCGTWTKAEGGIFTSPTNPSKYPDPRECI 68
 QY 70 YIIIEAPROCIIEFYDEKYSIEPSWECKFDHIEVRDGPFGFSPIIGRFGCGQNPPVTKSS 129
 DB 69 YIIIEAPROCIIEFYDEKYSIEPSWECKFDHIEVRDGPFGFSPIIGRFGCGQNPPVTKSS 128
 QY 130 GRFLWIKFPADGELSMGFSARYNFTF 156
 DB 129 GRFLWIKFPADGELSMGFSARYNFTF 155

```

RESULT 7
Q8ND78      PRELIMINARY; PRT; 500 AA.
ID 08ND78;
AC 08ND78;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, last annotation update)
DE Hypothetical protein (Fragment).
GN DKFZ547B197.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Ansgore W., Winkner U., Mewes H.W., Weil B., Wiemann S.;
RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AL84354; CAD39019.1; -.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR002172; LDL_receptor_A.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00057; ldl_recept_a; 1.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00192; LDLa; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS50068; LDLRA_2; 1.
KM Hypothetical protein.
FT NON_TER
SQ SEQUENCE 500 AA; 56708 MW; 313628B650C49F08 CRC64;

Query Match
Best Local Similarity 80.0%; Score 690; DB 4; Length 500;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 ETOKSVQCGTWTAKHAGGIFTSPNYSKYPPDRECIYIIIEAPROCIIEFYDEKYSIEPS 93
DB 1 ETOKSVQCGTWTAKHAGGIFTSPNYSKYPPDRECIYIIIEAPROCIIEFYDEKYSIEPS 60

QY 94 WECKFDHIEVRDGPFGSPITIGRFGCGQNPVYKSSGRFLWIKFFADGELSMGFSARVN 153
DB 61 WECKFDHIEVRDGPFGSPITIGRFGCGQNPVYKSSGRFLWIKFFADGELSMGFSARVN 120

QY 154 FTP 156
DB 121 FTP 123

RESULT 8
Q8NCM2      PRELIMINARY; PRT; 466 AA.
ID 08NCM2;
AC 08NCM2;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, last annotation update)
DE Hypothetical Microbodies C-terminal targeting signal/LDL-receptor
DE class A (Fragment).
GN NETO1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team.
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RT Nature 420:563-573 (2002).
DR EMBL; AK032510; BAC27902.1; -.

```

```

DR MGD; MG1:2180216; Netol.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR002172; LDL_receptor_A.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00057; ldl_recept_a; 1.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00192; LDLa; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS50068; LDLRA_2; 1.
KM Hypothetical protein.
FT NON_TER
SQ SEQUENCE 466 AA; 52840 MW; FE4AEA037EB344D9 CRC64;

Query Match
Best Local Similarity 57.3%; Score 494; DB 11; Length 466;
Matches 88; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 68 CYIIIEAPROCIIEFYDEKYSIEPSWECKFDHIEVRDGPFGSPITIGRFGCGQNPVYK 127
DB 1 CYIIIEAPROCIIEFYDEKYSIEPSWECKFDHIEVRDGPFGSPITIGRFGCGQNPVYK 60

QY 128 SSGRFLWIKFFADGELSMGFSARVNFTP 156
DB 61 SSGRFLWIKFFADGELSMGFSARVNFTP 89

RESULT 9
Q8NC67      PRELIMINARY; PRT; 525 AA.
ID 08NC67;
AC 08NC67;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, last annotation update)
DE Hypothetical protein FLJ90456.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA Masuko Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Niimoriya K.;
RT "NEBO human cDNA sequencing project.";
RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AK074937; BAC11303.1; -.
DR Genew; HGNC:14644; NETO2.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR002172; LDL_receptor_A.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00057; ldl_recept_a; 1.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00192; LDLa; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS50068; LDLRA_2; 1.
KM Hypothetical protein.
SQ SEQUENCE 525 AA; 59392 MW; EA6F98C3A88220EA CRC64;

Query Match
Best Local Similarity 54.5%; Score 470; DB 4; Length 525;
Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;

QY 6 SVLHT--VASLIILHSGATKKGTQKTSFQKSVQCGTWTAKHAGGIFTSPNYSKYTP 63
DB 8 SVLKVLITLVAVGIAVQKTDQGNIGIKHLPATQCGIWTSTNGHFPASPNYDSDYP 67

QY 64 PDRECIYIIIEAPROCIIEFYDEKYSIEPSWECKFDHIEVRDGPFGSPITIGRFGCGQNP 123
DB 68 PDRECIYIIIEAPROCIIEFYDEKYSIEPSWECKFDHIEVRDGPFGSPITIGRFGCGQNP 127

```

QY 124 PVKSGRFLWIKFFADGELSMGFSARYNFTP 156
 DB 128 PLIRSTGRFMWIKFSSDEBLEGLGFRAKYSFIP 160

RESULT 10

QY2381 PRELIMINARY; PRT; 491 AA.
 AC Q72381;
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DE Hypothetical protein DKFZp686N19198 (Fragment).
 GN DKFZp686N19198.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Human fetal brain;
 RA Blocher H., Boecher M., Mewes H.W., Weil B., Amid C., Oanger A.,
 RU Fodor G., Han M., Wiemann S.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BX538063; CAD97994.1; -.
 KW Hypothetical protein.
 FT NON TER
 SQ SEQUENCE 491 AA; 55641 MW; 3608CPA37BEIC6AC CRC64;

Query Match 54.3%; Score 468; DB 4; Length 491;
 Best Local Similarity 65.5%; Pred. No. 3,1e-43;
 Matches 78; Conservative 21; Mismatches 20; Indels 0; Gaps 0;

QY 38 SVQCGTAKAEGGIFSPNPKYPPDRECIYITAAAPROCIELYPDEKYSIEPSWECK 97
 DB 8 ATQCGIWRISNGGHFASPNYPDSYPPKCEIYITAEAPRQRIELTPEDEHYIIPSECR 67
 QY 98 FDHLEVRDGPFGFSPITIGRFGGQONPPYIKSSGRFLWIKFFADGELSMGFSARYNFTP 156
 DB 68 FDHLEVRDGPFGFSPITIGRFGGQONPPYIKSSGRFLWIKFFADGELSMGFSARYNFTP 126

RESULT 11

QY2381 PRELIMINARY; PRT; 444 AA.
 AC Q8C408;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DE Hypothetical LDL-receptor class A (Fragment).
 GN NETO2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RA MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium;
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573 (2002).
 DR EMBL; AK081462; BAC38225.1; -.
 DR MGD; MGI:1921763; NetO2.
 DR InterPro; IPR000859; CUB.
 DR InterPro; IPR002172; LDL_receptor_A.
 DR Pfam; PF00431; CUB; 2.
 DR Pfam; PF00057; ldl_recept_a; 1.
 DR SMART; SM00042; CUB; 2.
 DR SMART; SM00192; LDLa; 1.
 DR PROSITE; PS01180; CUB; 2.

DR PROSITE; PSS0068; LDLA_2; 1.
 KW Hypothetical protein.
 FT NON TER
 SQ SEQUENCE 444 AA; 49350 MW; 4CB0A5C93C52657E CRC64;

Query Match 53.9%; Score 465; DB 11; Length 444;
 Best Local Similarity 53.7%; Pred. No. 6e-43;
 Matches 79; Conservative 31; Mismatches 37; Indels 0; Gaps 0;

QY 10 IVASLIILHSGATKGTGKQTSTETQSKSVQCGTWTAKAEGGIFSPNPKYPPDRECI 69
 DB 14 LITVLVAGIAVAQKTQCGQNGICHIPTQCGIWRISNGGHFASPNYPDSYPPKCEI 146
 QY 70 YITAEAPROCIELYPDEKYSIEPSWECKFDHLEVRDGPFGFSPITIGRFGGQONPPYIKSS 129
 DB 147 YITAEAPRORIELTPEDEHYIIPSECRFDHLEVRDGPFGFSPITIGRFGGQONPPYIKSS 206
 QY 130 GRFLWIKFFADGELSMGFSARYNFTP 156
 DB 207 GRFLWIKFSSDEBLEGLGFRAKYSFIP 233

RESULT 12

QY2381 PRELIMINARY; PRT; 525 AA.
 AC Q8BN26;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DE Hypothetical LDL-receptor class A.
 GN NETO2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Body;
 RA MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium;
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573 (2002).
 DR EMBL; AK083512; BAC38938.1; -.
 DR MGD; MGI:1921763; NetO2.
 DR InterPro; IPR000859; CUB.
 DR InterPro; IPR002172; LDL_receptor_A.
 DR Pfam; PF00431; CUB; 2.
 DR Pfam; PF00057; ldl_recept_a; 1.
 DR SMART; SM00042; CUB; 2.
 DR SMART; SM00192; LDLa; 1.
 DR PROSITE; PS01180; CUB; 2.
 DR PROSITE; PSS0068; LDLA_2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 525 AA; 59367 MW; 3E0230B8F311EC5CE CRC64;

Query Match 53.9%; Score 465; DB 11; Length 525;
 Best Local Similarity 53.7%; Pred. No. 7.3e-43;
 Matches 79; Conservative 31; Mismatches 37; Indels 0; Gaps 0;

QY 10 IVASLIILHSGATKGTGKQTSTETQSKSVQCGTWTAKAEGGIFSPNPKYPPDRECI 69
 DB 14 LITVLVAGIAVAQKTQCGQNGICHIPTQCGIWRISNGGHFASPNYPDSYPPKCEI 73
 QY 70 YITAEAPROCIELYPDEKYSIEPSWECKFDHLEVRDGPFGFSPITIGRFGGQONPPYIKSS 129
 DB 74 YITAEAPRORIELTPEDEHYIIPSECRFDHLEVRDGPFGFSPITIGRFGGQONPPYIKSS 133
 QY 130 GRFLWIKFFADGELSMGFSARYNFTP 156
 DB 134 GRFLWIKFSSDEBLEGLGFRAKYSFIP 160

RESULT 13

```

ID 08ND51 PRELIMINARY; PRT; 451 AA.
AC 08ND51;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKFZP547L1153.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxId=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RL Bloecker H., Boecher M., Brandt P., Mewes H.W., Weil B., Wiemann S.;
Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AL834404; CAD39066.1; -.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR002172; LDL_receptor_A.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00057; LDL_recept_a; 1.
DR SMART; SM00042; CUB; 2.
DR SMART; SM01192; LDLa; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS50068; LDLRA_2; 1.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 451 AA; 51342 MW; 7BAD69517903B863 CRC64;

Query Match
Best Local Similarity 39.6%; Score 341; DB 4; Length 451;
Matches 59; Conservative 16; Mismatches 11; Indels 0; Gaps 0;

QY 71 IIEAAPROCIIEYFDEKYSIEPSWECKPHIEVRDGPSPFPIGFCGQNPVTKSSG 130
DB 1 IIEAAPROCIIEYFDEKYSIEPSWECKPHIEVRDGPSPFPIGFCGQNPVTKSSG 60
QY 131 RFLMIKFADGELSMGFSARYNFTP 156
DB 61 RFLMIKFSSDEHLEGLGFRAKSFRIP 86

```

RESULT 14

```

ID 061849 PRELIMINARY; PRT; 321 AA.
AC 061849;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE K03E5.1 protein.
GN K03E5.1
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodirinae; Caenorhabditis.
OC NCBI_TaxId=6239;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RL MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kerstan J., Kirtlen J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurry A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R.,
RA Smalton N., Smith A., Sonnenhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinlock L., Wilkinson-Sproat J., Woulman P.;

```

RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.";

RL Nature 368:32-38(1994).

RN (2)

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Becker M., Graves T., Fronick B.;

RT "The sequence of C. elegans cosmid K03E5.";

RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

RN (3)

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Waterston R.;

RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.

DR EMBL; AF067619; AAC17565.1; -.

DR PIR; T33161; T33161.

DR WormSep; K03E5.1; CE18007.

DR InterPro; IPR000859; CUB.

DR Pfam; PF00431; CUB; 2.

DR SMART; SM00042; CUB; 2.

DR PROSITE; PS01180; CUB; 2.

SQ SEQUENCE 321 AA; 36995 MW; 1E2BDDEFA70C3E36E CRC64;

Query Match
Best Local Similarity 30.2%; Score 260; DB 5; Length 321;
Matches 50; Conservative 20; Mismatches 37; Indels 8; Gaps 2;

QY 52 IFTSNVSKYPPEDDECIYIIEAAPROCIIEYFDEKYSIEPSM-----ECKPHIEVR 104
DB 75 IFTSNVSKYPPEDDECIYIIEAAPROCIIEYFDEKYSIEPSM-----ECKPHIEVR 134
QY 105 DGPFGFSPPIGRFCGQNP-VIKSGRFLMIKFADGELSMGFSARYNFTP 158
DB 135 DGPFGFSPPIGRFCGQNP-VIKSGRFLMIKFADGELSMGFSARYNFTP 189

RESULT 15

```

ID 09VYC7 PRELIMINARY; PRT; 677 AA.
AC 09VYC7; Q882J8;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE CG32635 protein (GH11189p).
GN CG32635 OR CG12727 OR CG15751 OR CG15752.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriidae; Drosophilidae; Drosophila.
OC NCBI_TaxId=7227;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RL MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananthakrishnan P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortan J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan R.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-U., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Borzhan M.R., Bouck J., Brockstein P., Brotter P.,
RA Burdick K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Dew I., Dietz S.M.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferrel R., Ferrel R., Ferrel R.,
RA Rosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 18:16:31 ; Search time 17 Seconds

(without alignments)
483.946 Million cell updates/sec

Title: US-09-887-593a-2
Perfect score: 862
Sequence: 1 MHGRSVLHIVASLLIHLIS.....ADGHELSMGFSARNVTPGX 158

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	232.5	27.0	914	1	NRPI_CHICK
2	227	26.3	707	1	BMPI_XENLA
3	226	26.2	931	1	NRPI_MOUSE
4	225	26.1	925	1	NRPI_MOUSE
5	224	26.0	931	1	NRPI_MOUSE
6	224	26.0	931	1	NRPI_MOUSE
7	221	25.6	923	1	NRPI_MOUSE
8	220.5	25.6	986	1	BMPI_MOUSE
9	220.5	25.6	1022	1	TUD_BRARE
10	217.5	25.2	1057	1	TUD_BRARE
11	217	25.2	922	1	NRPI_MOUSE
12	216	25.1	923	1	NRPI_MOUSE
13	215	24.9	639	1	BMPI_MOUSE
14	201.5	23.4	449	1	PCOI_MOUSE
15	198.5	23.0	928	1	NRPI_MOUSE
16	194.5	22.6	468	1	PCOI_MOUSE
17	194.5	22.6	468	1	PCOI_MOUSE
18	186	21.6	3564	1	CSMI_MOUSE
19	184.5	21.4	3565	1	CSMI_MOUSE
20	182	21.1	3487	1	CSMI_MOUSE
21	178.5	20.7	599	1	CRAR_MOUSE
22	173.5	20.1	597	1	CRAR_MOUSE
23	163.5	19.0	2796	1	CSM3_MOUSE
24	163.5	19.0	3670	1	CSM3_MOUSE
25	162.5	18.9	704	1	CRAR_MOUSE
26	161.5	18.7	616	1	SPAN_MOUSE
27	160.5	18.6	616	1	SPAN_MOUSE
28	152.5	17.7	705	1	CRAR_MOUSE
29	151	17.5	275	1	TS66_MOUSE
30	151	17.5	277	1	TS66_MOUSE
31	150	17.4	310	1	ASTL_MOUSE
32	149.5	17.3	855	1	ST14_MOUSE
33	148	17.2	276	1	TS66_MOUSE

34	142.5	16.5	514	1	UVS2_XENLA
35	136.5	15.8	855	1	ST14_MOUSE
36	129.5	15.0	1428	1	ATRN_MOUSE
37	127.5	14.8	1429	1	ATRN_MOUSE
38	116	13.5	1034	1	ENTK_PIG
39	116	13.5	1069	1	ENTK_MOUSE
40	115	13.3	462	1	KRM2_MOUSE
41	114	13.2	1035	1	ENTK_MOUSE
42	111	12.9	695	1	CASP_MOUSE
43	110	12.8	686	1	MAS2_MOUSE
44	107.5	12.5	133	1	AMN_MOUSE
45	107.5	12.5	688	1	CIS_MOUSE

ALIGNMENTS

RESULT 1
NRPI_CHICK
ID NRPI_CHICK STANDARD; PRT; 914 AA.
AC P79795;
DT 01-NOV-1997 (Rel. 35, Created)
DI 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neupilin-1 precursor (A5 protein).
GN NRPI OR NRPI.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn; TISSUE=Embryonic brain;
RX MEDLINE=95324761; PubMed=7601310;
RA Takagi S., Kasuya Y., Shimizu M., Matsura T., Tsuboi M., Kawakami A.,
Fujisawa H.;
RT "Expression of a cell adhesion molecule, neupilin, in the
RT developing chick nervous system.";
RL Dev. Biol. 170:207-222(1995).
CC -!- FUNCTION: Receptor involved in the development of the
CC cardiovascular system, in angiogenesis, in the formation of
CC certain neuronal circuits and in organogenesis outside the nervous
CC system. It mediates the chemorepulsive activity of semaphorins (By
CC similarity). Seems to have calcium-independent cell adhesion
CC properties.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Developing nervous system; optic tectum
CC (layers D and E of SGRS), amacrine cells of retina, neurites of
CC dorsal root ganglia. Also expressed in nonneuronal cells, e.g.
CC blood vessels in the entire embryo.
CC -!- SIMILARITY: Belongs to the neupilin family.
CC -!- SIMILARITY: Contains 2 CUB domains.
CC -!- SIMILARITY: Contains 2 F5/8 type C domains.
CC -!- SIMILARITY: Contains 1 MAM domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D45416; BAA08256.1; -.
CC HSRP: P12259; ICZT.
CC InterPro: IPR000859; CUB.
CC InterPro: IPR000421; FAS8_C.
CC InterPro: IPR008979; Gal_bind_like.
CC InterPro: IPR000998; MAM_domain.
CC Pfam: PF00431; CUB; 2.
CC Pfam: PF00754; F5_P8_type_C; 2.
CC Pfam: PF00629; MAM; 1.

```

DR PRINTS: PR00020; MAMDOMAIN.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00231; FAS8C; 2.
DR SMART; SM00137; MAM; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS01285; FAS8C_1; 2.
DR PROSITE; PS01286; FAS8C_2; 2.
DR PROSITE; PS00022; FAS8C_3; 2.
DR PROSITE; PS00740; MAM_1; 1.
DR PROSITE; PS00060; MAM_2; 1.
DR Angiogenesis; Transmembrane; Glycoprotein; Neurone; Signal; Repeat;
DR Receptor; Cell adhesion.
KW SIGNAL
FT CHAIN 1 18
FT DOMAIN 19 914 NEUROBLIN-1
FT TRASMEM 20 847 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 848 870 POTENTIAL.
FT DOMAIN 871 914 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 139 CUB 1.
FT DOMAIN 145 263 CUB 2.
FT DOMAIN 273 422 FS/8 TYPE C 1.
FT DOMAIN 429 581 FS/8 TYPE C 2.
FT DOMAIN 636 801 MAM.
FT DISULFID 25 52 PROBABLE.
FT DISULFID 80 102 PROBABLE.
FT DISULFID 145 171 PROBABLE.
FT DISULFID 204 226 PROBABLE.
FT DISULFID 273 422 BY SIMILARITY.
FT DISULFID 429 581 BY SIMILARITY.
SQ SEQUENCE 914 AA; 102480 MW; DD2EED6F0CB68C CRC64;

Query Match 27.0%; Score 232.5; DB 1; Length 914;
Best Local Similarity 35.3%; Pred. No. 4e-16; Indels 19; Gaps 4;
Matches 54; Conservative 25; Mismatches 55;

QY 15 IILHSGATKTKGTEKOTSETOKSVOCGWTAKABGIFLSPNPKYPPDRRCIYIA 74
DB 5 LFLHCAALT-----FTLSALAKSDKCGPTIKILSPGYLTPGYPSQSYHPSQKCEMLIO 58
QY 75 -APROCIETYDEPKYIEPSMECKPHIEVRDPPFSPILGFCQGNPPVYKSGRFL 133
DB 59 PEYQQRIMNFNFHFLIEDR-DCKYDVEYIDGNAEGRLMKRYCGKIAPPLVSSGPYI 117
QY 134 WIKFPADGELESWGFSAFY-----NFT 155
DB 118 FIKFVSDYETHGAGFIRYEVFRKGPCCSNFT 150

DB
BML1_XENIA STANDARD; PRT; 707 AA.
ID BML1_XENIA
AC P98070;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Bone morphogenetic protein 1 precursor (EC 3.4.24.-) (BMP-1).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodidae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo; PubMed=8262384;
RA Maeno M., Xue Y., Wood T.I., Ong R.C., Kung H.F.;
RT "Cloning and expression of cDNA encoding Xenopus laevis bone
morphogenetic protein-1 during early embryonic development.";
RT Gene 134:257-261(1993).
RL -1- FUNCTION: Involved in pattern formation in gastrula and later
differentiation of developing organs.
CC -1- DEVELOPMENTAL STAGE: Blastula, early gastrula and hatched
caddoles; little or no expression in morula and late gastrula.
CC -1- SIMILARITY: Belongs to peptidase family M12A.

```

```

CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 3 CUB domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J12249; AAA16313.1; -.
DR PIR; JC2218; JC2218.
DR HSP; P00736; IAPQ.
DR MEROPS; M12.005; -.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR006026; Peptidase_M.
DR InterPro; IPR001506; Peptidase_M12A.
DR Pfam; PF01400; Ascatin; 1.
DR Pfam; PF00431; CUB; 3.
DR Pfam; PF00008; EGF; 3.
DR PRINTS; PR00480; ASTACIN.
DR SMART; SM00042; CUB; 3.
DR SMART; SM00179; BGF_CA; 1.
DR SMART; SM00235; ZMGC; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS01180; CUB; 3.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR Growth factor; Cytokine; Repeat; Osteogenesis; Chondrogenesis;
KW Hydrolyase; Protease; Metalloprotease; EGF-like domain; Zinc; Calcium;
KW Signal.
FT SIGNAL 1 2
FT PROPEP 83 83
FT CHAIN 84 707
FT DOMAIN 84 284 BONE MORPHOGENETIC PROTEIN 1.
FT DOMAIN 285 397 METALLOPROTEASE.
FT DOMAIN 398 509 CUB 1.
FT DOMAIN 510 551 CUB 2.
FT DOMAIN 554 666 CUB 3.
FT METAL 176 176 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT SITE 177 177 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 180 180 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 186 186 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 146 149 BY SIMILARITY.
FT DISULFID 514 526 BY SIMILARITY.
FT DISULFID 522 535 BY SIMILARITY.
FT DISULFID 537 550 BY SIMILARITY.
FT CARBOHYD 62 62 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 105 105 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 326 326 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 562 562 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 707 AA; 80673 MW; 1B6980D716DC98BD CRC64;

Query Match 26.3%; Score 227; DB 1; Length 707;
Best Local Similarity 40.2%; Pred. No. 1.1e-15;
Matches 45; Conservative 18; Mismatches 47; Indels 2; Gaps 2;

QY 41 CGWTAKABGIFLSPNPKYPPDRRCIYIAAPROCIETYDEPKYIEPSMECKPH 100
DB 398 CGGEVKKDSCHI-QSPNYPDPDYRPNKACVWKLVSSEGFHGISF-QSEIERHDSCAVDY 455
QY 101 IEVRDPPFSGSPILGRCQGNPPVYKSGRFLMIKFPADGELESWGFSAFY 152
DB 456 IEIRDSSETSPLVGRFCGYDKDPDIKSTNQILKIFVSDGSIKAGPSLANY 507

```



```

RESULT 3
NRP2_MOUSE
ID NRP2_MOUSE STANDARD, PRT, 931 AA.
AC Q35375; Q35373; Q35374; Q35376; Q35377; Q35378;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neuphilin-2 precursor (Vascular endothelial cell growth factor 165
  receptor 2).
GN NRP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A. (ISOFORMS A0, A17, A22, A5; B0 AND B5).
RC STRAIN=BALB/c;
RA Chem.H.; 97470888; PubMed=9311348;
RA Chen H., Chedotal A., He Z.-G., Goodman C.S., Tessier-Lavigne M.;
RT "Neuphilin-2, a novel member of the neuphilin family, is a high
RT affinity receptor for the semaphorins Sema E and Sema IV but not Sema
RT III.";
RL Neuron 19:547-559 (1997)
CC -1- FUNCTION: High affinity receptor for semaphorins 3C, 3F, VEGF-165
CC and VEGF-145 isoforms of VEGF, and the PLGF-2 isoform of PGF.
CC -1- SUBUNIT: Neuphilin-2 probably forms a heteromeric complex with
CC neuphilin-1 in order to be a functional semaphorin E receptor.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=6;
CC Name=A22;
CC IsoId=Q35375-1; Sequence=Displayed;
CC Name=A0;
CC IsoId=Q35375-2; Sequence=VSP_004344;
CC Name=A5;
CC IsoId=Q35375-3; Sequence=VSP_004345;
CC Name=A17;
CC IsoId=Q35375-4; Sequence=VSP_004343;
CC Name=B0;
CC IsoId=Q35375-5; Sequence=VSP_004346;
CC Name=B5;
CC IsoId=Q35375-6; Sequence=VSP_004347;
CC -1- TISSUE SPECIFICITY: Expressed in developing CNS, PNS and in some
CC nonneural tissues including limb buds, developing bones, muscles,
CC intestinal epithelium, kidney, lung and submandibular gland.
CC -1- DEVELOPMENTAL STAGE: The expression pattern is very dynamic and is
CC developmentally regulated.
CC -1- SIMILARITY: Belongs to the neuphilin family.
CC -1- SIMILARITY: Contains 2 CUB domains.
CC -1- SIMILARITY: Contains 2 F5/8 type C domains.
CC -1- SIMILARITY: Contains 1 MAM domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF022856; AAC53379.1; -
DR EMBL; AF022854; AAC53377.1; -
DR EMBL; AF022855; AAC53378.1; -
DR EMBL; AF022857; AAC53380.1; -
DR EMBL; AF022858; AAC53381.1; -
DR EMBL; AF022861; AAC53382.1; -
DR HSSP; P12259; 1CZT.
DR MGD; MGI:1100492; NrP2.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR000421; FAS5_C.
DR InterPro; IPR008979; Gal_Bind_Like.

```

DR InterPro: IPR000398; MAM_domain.
DR Pfam: PF00431; CUB; 2.
DR Pfam: PF00754; F5_F8_type_C; 2.
DR Pfam: PF00629; MAM; 1.
DR PRINTS: PR00020; MAMDOMAIN.
DR SMART: SM00043; CUB; 2.
DR SMART: SM00231; FAS5C; 2.
DR SMART: SM00137; MAM; 1.
DR PROSITE: PS01180; CUB; 2.
DR PROSITE: PS01285; FAS5C_1; 2.
DR PROSITE: PS01286; FAS5C_2; 2.
DR PROSITE: PS50022; FAS5C_3; 2.
DR PROSITE: PS50060; MAM_2; 1.
KW Transmembrane; Glycoprotein; Neutrone; Signal; Repeat; Receptor;
KW Alternative splicing.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 931 NEUROPEPTILIN-2.
FT DOMAIN 21 864 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 865 889 POTENTIAL.
FT DOMAIN 890 931 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 28 142 CUB 1.
FT DOMAIN 149 267 CUB 2.
FT DOMAIN 277 427 F5/8 TYPE C 1.
FT DOMAIN 434 592 F5/8 TYPE C 2.
FT DOMAIN 642 802 MAM.
FT DOMAIN 838 845 POLY-SER.
FT DISULFID 28 55 BY SIMILARITY.
FT DISULFID 83 105 BY SIMILARITY.
FT DISULFID 149 175 BY SIMILARITY.
FT DISULFID 208 230 BY SIMILARITY.
FT DISULFID 277 427 BY SIMILARITY.
FT DISULFID 434 592 BY SIMILARITY.
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 839 839 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VASPPPLIC 809 813 Missing (in isoform A17).
FT VASPPPLIC 809 830 /FtId=VSP_004343.
FT VASPPPLIC 809 830 Missing (in isoform A0).
FT VASPPPLIC 814 830 Missing (in isoform A0).
FT VASPPPLIC 814 830 Missing (in isoform A5).
FT VASPPPLIC 810 931 /FtId=VSP_004345.
FT VASPPPLIC 810 931 EDPEVDVLPETHGSGEGEIDEIDEYEGDMSSSSSTSGADP
FT FT SSGEKEMVLYTIDPILITIIAMSSLSGLVIGATCAGLLCYCT
FT FT CSYSGLSRSCITLLENVPELYDGLKHKVYKINHOCCSEA
FT FT -> GLTPGTEPTVDIVPQPIPAWYIYVMAAGAVIYLAS
FT FT VIALVLIHYHFRYAAKTDHSTIYKTSHTYNGAPLAVEPT
FT FT LTIKEORSGHC (in isoform B0).
FT FT /FtId=VSP_004346.
FT FT VDIPEITGGEIEDEIDEYEGDMSSSSSTSGADPSSGK
FT FT ESMWLYTIDPILITIIAMSSLSGLVIGATCAGLLCYCTCSY
FT FT GLASRSCITLLENVPELYDGLKHKVYKINHOCCSEA -> G
FT FT GLTPGTEPTVDIVPQPIPAWYIYVMAAGAVIYLASVYL
FT FT ALVLIHYHFRYAAKTDHSTIYKTSHTYNGAPLAVEPTLTI
FT FT KLEORSGHC (in isoform B5).
FT FT /FtId=VSP_004347.
FT FT G->I (in REF. 1; AAC53380/AAC53381).
FT SEQUENCE 786 786
FT 931 AA; 104558 MW; 76P2443P411D2F63 CRC64;
Query Match 26.2%; Score 226; DB 1; Length 931;
Best Local Similarity 38.1%; Pred. No. 1.9e-15;
Matches 51; Conservative 18; Mismatches 51; Indels 14; Gaps 3;

Db 141 IFKTGSBDSCKNFT 154

RESULT 4

NRP2 RAT STANDARD; PRT: 925 AA.

AC 035276;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Neutropilin-2 precursor (Vascular endothelial cell growth factor 165 receptor 2).

GN NRP2.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_Taxid=101116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Sprague-Dawley;

RX MEDLINE=97433085; PubMed=9288754;

RA Kolodkin A.L., Levengood D.V., Rowe E.G., Tai Y.-T., Giger R.J., Ginry D.D.;

RT "Neutropilin is a semaphorin III receptor.";

RU Cell 90:753-762(1997).

CC -1- FUNCTION: High affinity receptor for semaphorins 3C, 3F, VEGF-165 and VEGF-145 isoforms of VEGF, and the PLGF-2 isoform of PGF.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- TISSUE SPECIFICITY: Found in certain neuronal populations of the CNS and in other nonneuronal tissues including mesenchymal tissue lining in the ribs.

CC -1- SIMILARITY: Belongs to the neutropilin family.

CC -1- SIMILARITY: Contains 2 CUB domains.

CC -1- SIMILARITY: Contains 2 F5/8 type C domains.

CC -1- SIMILARITY: Contains 1 MAM domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

CC EMBL: AF016297; AAC5338.1; -.

CC HSSP: P12259; ICZT.

CC InterPro: IPR000859; CUB.

CC InterPro: IPR000421; FAS8.C.

CC InterPro: IPR008979; GalBind like.

CC InterPro: IPR000998; MAM_domain.

CC Pfam: PF00431; CUB_2.

CC Pfam: PF00754; F5_F8_type_C_2.

CC Pfam: PF00629; MAM_1.

CC PRINTS: PRO0020; MAMDOMAIN.

CC SMART: SM00042; CUB_2.

CC SMART: SM00231; FAS8C_2.

CC SMART: SM00137; MAM_1.

CC PROSITE: PS01180; CUB_2.

CC PROSITE: PS01285; FAS8C_1; 2.

CC PROSITE: PS01286; FAS8C_2; 2.

CC PROSITE: PS50022; FAS8C_3; 2.

CC PROSITE: PS50060; MAM_2; 1.

CC Transmembrane; Glycoprotein; Neutrine; Signal; Repeat; Receptor.

KW SIGNA1.

FT CHAIN 1 22

FT DOMAIN 23 925 NEUTROPILIN-2.

FT TRANSMEM 859 883 EXTRACELLULAR (POTENTIAL).

FT DOMAIN 884 925 POTENTIAL.

FT DOMAIN 28 142 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 149 267 CUB 1.

FT DOMAIN 277 427 CUB 2.

FT DOMAIN 434 592 F5/8 TYPE C 1.

FT DOMAIN 642 802 F5/8 TYPE C 2.

FT MAM.

FT DISULFID 28 55 BY SIMILARITY.

FT DISULFID 83 105 BY SIMILARITY.

FT DISULFID 149 175 BY SIMILARITY.

FT DISULFID 208 230 BY SIMILARITY.

FT DISULFID 277 427 BY SIMILARITY.

FT DISULFID 434 592 BY SIMILARITY.

FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 833 833 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 834 834 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 925 AA; 103896 MW; 3BF62903F644851C CRC64;

Query Match 26.1%; Score 225; DB 1; Length 925;

Best Local Similarity 37.3%; Pred. No. 2,4e-15;

Matches 50; Conservative 19; Mismatches 51; Indels 14; Gaps 3;

Db 153 -----NFT 155

141 IFKTGSBDSCKNFT 154

QY 35 TQSVQCGTWTGKAGGIFTSFNPYSKTPPPDRECIYITIA-APROCIETLPDEKSIERS 93

Db 22 SQQDPCCGRUNSKDAGYITSPGYQDYPHONCEWVYAPENOKIVTFNPFHEIE-K 80

QY 94 WECKEDHIEVRDPPGFPPIIGRFGQONPPYIKSGRFLMKFPADGELSEMGFSARY- 152

Db 81 HDCKYDFLIRGDSBSADLKGKHCNTAPITLISGSVLYIKFTSDYARQAGFSLRYE 140

QY 153 -----NFT 155

141 IFKTGSBDSCKNFT 154

RESULT 5

NRP2 HUMAN STANDARD; PRT: 931 AA.

ID NR22 HUMAN

AC 060462; O14820; O14821;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Neutropilin-2 precursor (Vascular endothelial cell growth factor 165 receptor 2).

GN NRP2 OR VEGF165R2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM A0 AND A17).

RP MEDLINE=97470886; PubMed=9331348;

RA Chen H., Chedotal A., He Z.-G., Goodman C.S., Tessier-Lavigne M., Chen H., Chedotal A., He Z.-G., Goodman C.S., Tessier-Lavigne M.;

RT "Neutropilin-2, a novel member of the neutropilin family, is a high affinity receptor for the semaphorins Sema E and Sema IV but not Sema III.";

RT Neuron 19:547-559(1997).

RU [2]

RP SEQUENCE FROM N.A. (ISOFORM A22).

RP TISSUE=Breast;

RC MEDLINE=98188099; PubMed=9529250;

RA Soker S., Takashima S., Miao H.-Q., Neufeld G., Klagsbrun M.;

RT "Neutropilin-1 is expressed by endothelial and tumor cells as an isoform-specific receptor for vascular endothelial growth factor.";

RT Cell 92:735-745(1998).

RN [3]

RP CHARACTERIZATION.

RP MEDLINE=20309748; PubMed=10748121;

RA Guzman-Polcorak Z., Cohen T., Herzog Y., Neufeld G.;

RT "Neutropilin-2 and neutropilin-1 are receptors of the 165-amino acid form of vascular endothelial growth factor (VEGF) and of placenta growth factor-2, but only neutropilin-2 functions as a receptor for the 145-amino acid form of VEGF.";

RT J. Biol. Chem. 275:18040-18045(2000).

CC -1- FUNCTION: High affinity receptor for semaphorins 3C, 3F, VEGF-165 and VEGF-145 isoforms of VEGF, and the PLGF-2 isoform of PGF.

CC -1- SUBUNIT: Neutropilin-2 probably forms a heteromeric complex with neutropilin-1 in order to be a functional semaphorin 3C receptor.

```

CC -|- SUBCELLULAR LOCATION: Type I membrane protein.
CC -|- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=3;
CC Name=A22;
CC IsoId=O60462-1; Sequence=Displayed;
CC Name=A0;
CC IsoId=O60462-2; Sequence=VSP_004342;
CC Name=A17;
CC IsoId=O60462-3; Sequence=VSP_004341;
CC -|- SIMILARITY: Belongs to the neuropilin family.
CC -|- SIMILARITY: Contains 2 CUB domains.
CC -|- SIMILARITY: Contains 2 F5/8 type C domains.
CC -|- SIMILARITY: Contains 1 MAM domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF022859; AAC51788.1; -
CC EMBL; AF022860; AAC51789.1; -
CC EMBL; AF016098; AAC12922.1; -
CC HSSP; P12259; 1CCT
CC Genew; HGNC:8005; NRP2.
CC MIM; 602070; -
CC DR GO:0005624; C:membrane fraction; TAS.
CC DR GO:0004872; F:receptor activity; TAS.
CC DR GO:0005021; F:vascular endothelial growth factor receptor. . .; TAS.
CC DR GO:0007411; P:axon guidance; TAS.
CC InterPro: IPR000859; CUB.
CC InterPro: IPR000421; FAS8_C.
CC InterPro: IPR008979; Gal Bind like.
CC InterPro: IPR000998; MAM_domain.
CC Pfam; PF00754; CUB; 2.
CC Pfam; PF00431; CUB; 2.
CC Pfam; PF00629; MAM; 1.
CC PRINTS; PR00020; MAMDOMAIN.
CC SMART; SMO0042; CUB; 2.
CC SMART; SMO0231; FAS8C; 2.
CC SMART; SMO0137; MAM; 1.
CC PROSITE; PS01180; CUB; 2.
CC PROSITE; PS01285; FAS8C_1; 2.
CC PROSITE; PS01286; FAS8C_2; 2.
CC PROSITE; PS50022; FAS8C_3; 2.
CC PROSITE; PS50060; MAM_2; 1.
CC Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
CC Alternative splicing.
CC SIGNAL 1 20
CC CHAIN 21 931
CC DOMAIN 21 864
CC TRANSEM 865 889
CC DOMAIN 890 931
CC DOMAIN 28 142
CC DOMAIN 149 267
CC DOMAIN 277 427
CC DOMAIN 434 592
CC DOMAIN 642 802
CC DOMAIN 671 674
CC DISULFID 28 55
CC DISULFID 83 105
CC DISULFID 149 175
CC DISULFID 208 230
CC DISULFID 277 427
CC CARBOHYD 434 592
CC CARBOHYD 152 152
CC CARBOHYD 157 157
CC CARBOHYD 629 629
CC CARBOHYD 839 839
CC VARSPIC 809 813

```

```

FT VARSPLIC 809 830 Missing (in isoform A0).
FT FT /FTId=VSP_004342.
FT CONFLICT 602 602 E -> K (IN REF. 1).
SQ SEQUENCE 931 AA; 104830 MW; 270CBAB69A0A797C CRC64;
Query Match 26.0%; Score 224; DB 1; Length 911;
Best local similarity 38.3%; Fred. No. 3.1e-15;
Matches 51; Conservative 16; Mismatches 52; Indels 14; Gaps 3;
QY 36 QKSVCGTWTHAEGIFTPSPYPSKYPDPRECIYIIEA-APROCIETLPDEKYSLEPSW 94
DB 23 QPDPCCGRSLNSKAGYITSFCYPODYSHQNCENTVAPEDNQXIVNPNFETIE-KH 81
QY 95 ECKEDHIEVRDGPFGFPDIIIRFCGQNPVYIKSSGRPLWIKFADGELESMGFSARY-- 152
DB 82 DCKDVFIRIDGDSRADLLKHGQNIAPPTIISGSMLYIKFTSDYARQGGFSLRYEI 141
QY 153 -----NFT 155
DB 142 FKTSSEDCSKNFT 154
RESULT 6
EMBL_MOUSE STANDARD; PRT; 991 AA.
AC P98063;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Bone morphogenetic protein 1 precursor (BMP-3.4.24.19) (BMP-1)
DE (Procollagen C-proteinase) (PCP) (Mammalian tollid protein) (mtId).
GN BMP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Embryo;
RX MEDLINE=94229342; PubMed=8174772;
RA Fukagawa M., Noboru S., Hogan B.L.M., Jones C.M.;
RT "Embryonic expression of mouse bone morphogenetic protein-1 (BMP-1),
RT which is related to the Drosophila dorsventral gene tollid and
RL encodes a putative astacin metalloendopeptidase."
CL Dev. Biol. 163:175-183(1994).
CC -|- FUNCTION: Cleaves the C-terminal propeptides of procollagen I, II
CC and III. Induces cartilage and bone formation.
CC -|- CATALYTIC ACTIVITY: Cleavage of the C-terminal propeptide at
CC Ala-I-Asp in type I and II procollagens and at Arg-I-Asp in type
CC III.
CC -|- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -|- ENZYME REGULATION: Activity is increased by the procollagen C-
CC endopeptidase enhancer protein.
CC -|- TISSUE SPECIFICITY: At high levels in embryonic maternal deciduum
CC and floor plate region of the neural tube. Less in developing
CC membranous and endochondral bone, submucosa of intestine, dermis
CC of skin and the mesenchyme of spleen and lung.
CC -|- SIMILARITY: Belongs to peptidase family M12A.
CC -|- SIMILARITY: Contains 2 EGF-like domains.
CC -|- SIMILARITY: Contains 5 CUB domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L24755; AAA37306.1; -
CC PTR: I49540; I49540.
DR HSSP: P00736; IAPQ.
DR MEROPS; M12.005; -.

```

DR MGD; MGI:88176; Bmp1.
 DR InterPro; IPR000152; ASX_hydroxyl_S.
 DR InterPro; IPR000859; CUB.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR006209; EGF_Like.
 DR InterPro; IPR006025; Pcdt_M_Zn_BS.
 DR InterPro; IPR006026; Peptidase_M2A.
 DR InterPro; IPR001506; Peptidase_M12A.
 DR Pfam; PF01400; Astacin; 1.
 DR Pfam; PF00431; CUB; 5.
 DR Pfam; PF00008; EGF; 2.
 DR PRINTS; PR00480; ASTACIN.
 DR SMART; SM00042; CUB; 5.
 DR SMART; SM00179; EGF_CA; 2.
 DR SMART; SM00235; ZMNC; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; 2.
 DR PROSITE; PS01180; CUB; 5.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS50026; EGF_3; 2.
 DR PROSITE; PS01187; EGF_CA; 2.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR Growth factor; Cytokine; Repeat; Osteogenesis; Chondrogenesis;
 KW Glycolase; Metalloprotease; EGF-like domain; Zinc; Calcium; Signal;
 KW Glycoprotein; Zymogen.
 FT SIGNAL 1 25
 FT PROPEP 26 125
 FT CHAIN 126 991
 FT DOMAIN 126 326
 FT DOMAIN 327 439
 FT DOMAIN 440 551
 FT DOMAIN 552 593
 FT DOMAIN 596 707
 FT DOMAIN 708 748
 FT DOMAIN 752 864
 FT METAL 865 981
 FT METAL 218 218
 FT ACT_SITE 219 219
 FT METAL 222 222
 FT METAL 228 228
 FT DISULFID 188 191
 FT DISULFID 327 353
 FT DISULFID 380 402
 FT DISULFID 440 466
 FT DISULFID 493 515
 FT DISULFID 556 577
 FT DISULFID 564 577
 FT DISULFID 579 592
 FT DISULFID 596 622
 FT DISULFID 649 671
 FT DISULFID 712 723
 FT DISULFID 719 732
 FT DISULFID 734 747
 FT CARBOHYD 96 96
 FT CARBOHYD 147 147
 FT CARBOHYD 337 337
 FT CARBOHYD 368 368
 FT CARBOHYD 604 604
 SQ SEQUENCE 991 AA; 111607 MW; 68A184783A0B89E CRC64;

Query Match 26.0%; Score 224; DB 1; Length 991;
 Best Local Similarity 33.6%; Pred. No. 3.3e-15;
 Matches 43; Conservative 27; Mismatches 42; Indels 16; Gaps 2;

QY 26 GTEKQTTSETQKSVQCGTWTAKBAGGIFTSPPNYSKYPPDRRCITYIIIAAPRCOTIELYFD 85
 DB 751 GCHKXTYMS-----GTTSPNWDKYPKSKKCTWMAISSTPGHRYKLTFFV 795
 QY 86 EKKSIEPSWCKRDHIEVDPGPGSPITIGRFGCOONPPVKKSGFRFWTKFPADELES 145
 DB 796 E-MDISQPCAVDHLVEFDGRDAKAPVLGRFCGSKKPEPVLATGNRMFLRFYSNDSVQR 854
 QY 146 MGFSARYN 153

DB 855 KGFGASHS 862
 |||
 ::
 RESULT 7
 ID NRPI_HUMAN STANDARD; PRT; 923 AA.
 AC 014786; 060461;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2003 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neupophilin-1 precursor (Vascular endothelial cell growth factor 165 receptor).
 DE receptor).
 GN NRPI OR NRP OR VEGF165R.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=97433084; PubMed=9288753;
 RA He Z., Tessier-Lavigne M.;
 RT "Neupophilin is a receptor for the axonal chemorepellent semaphorin III.";
 RT Cell 90:739-751(1997).
 RL [2]
 RN RP SEQUENCE FROM N.A. (ISOFORM 1), AND SEQUENCE OF 22-39.
 RC TISSUE=Brain;
 RX MEDLINE=98188099; PubMed=9529250;
 RA Soker S., Takashima S., Miao H.-Q., Neufeld G., Klagsbrun M.;
 RT "Neupophilin-1 is expressed by endothelial and tumor cells as an isoform-specific receptor for vascular endothelial growth factor.";
 RL Cell 92:735-745(1998).
 RP [3]
 RN RP SEQUENCE FROM N.A. (ISOFORM 2), AND SEQUENCE OF 22-31.
 RC TISSUE=Prostatic adenocarcinoma;
 RX MEDLINE=20183929; PubMed=1068880;
 RA Gagnon M.L., Bielebun D.R., Gschman Z., Miao H.-Q., Takashima S., Soker S., Klagsbrun M.;
 RT "Identification of a natural soluble neupophilin-1 that binds vascular endothelial growth factor: In vivo expression and anticancer activity.";
 RT Proc. Natl. Acad. Sci. U.S.A. 97:2573-2578(2000).
 RL [4]
 RN RP CHARACTERIZATION.
 RX MEDLINE=20309748; PubMed=10748121;
 RA Gluzman-Poltorak Z., Cohen T., Herzog Y., Neufeld G.;
 RT "Neupophilin-2 and neupophilin-1 are receptors for the 165-amino acid form of vascular endothelial growth factor (VEGF) and of placenta growth factor-2, but only neupophilin-2 functions as a receptor for the 145-amino acid form of VEGF.";
 RT J. Biol. Chem. 275:18040-18045(2000).
 RL [5]
 RN RP -1- FUNCTION: The membrane-bound isoform 1 is a receptor involved in the development of the cardiovascular system, in angiogenesis, in the formation of certain neuronal circuits and in organogenesis outside the nervous system. It mediates the chemorepellant activity of semaphorins. It binds to semaphorin 3A, The PlGF-2 isoform of PGF. The VEGF-165 isoform of VEGF and VEGF-B. Coexpression with KDR results in increased VEGF-165 binding to KDR as well as increased chemotaxis. It may regulate VEGF-induced angiogenesis.
 CC -1- FUNCTION: The soluble isoform 2 binds VEGF-165 and appears to inhibit its binding to cells. It may also induce apoptosis by sequestering VEGF-165. May bind as well various members of the semaphorin family. Its expression has an adverse effect on blood vessel number and integrity.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Isoform 2 is secreted.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=Membrane-bound;
 CC IsoId=O14786-1; Sequence=Displayed;
 CC Name=2; Synonyms=Soluble, SNRP1;

CC IsoId=014786-2; Sequence=VSP_004339, VSP_004340;
 CC -1- TISSUE SPECIFICITY: The expression of isoforms 1 and 2 does not
 CC seem to overlap. Isoform 1 is expressed by the blood vessels of
 CC different tissues. In the developing embryo it is found
 CC predominantly in the nervous system. In adult tissues, it is
 CC highly expressed in heart and placenta; moderately in lung, liver,
 CC skeletal muscle, kidney and pancreas; and low in adult brain.
 CC Isoform 2 is found in liver hepatocytes, kidney distal and
 CC proximal tubules.
 CC -1- SIMILARITY: Belongs to the neuropilin family.
 CC -1- SIMILARITY: Contains 2 CUB domains.
 CC -1- SIMILARITY: Contains 2 F5/8 type C domains.
 CC -1- SIMILARITY: Contains 1 MAM domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch)
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF018956; AAC51759.1; -
 CC EMBL; AF016050; AAC12921.1; -
 CC EMBL; AF145712; AAF4344.1; -
 CC PDB; 1KEX; 28-JAN-03.
 CC Genew; HGNC:8004; NRPL.
 CC MIM; 602069; -
 CC DR GO; GO:0005021; F:vascular endothelial growth factor receptor. . . ; TAS.
 CC DR GO; GO:0007411; P:axon guidance; TAS.
 CC DR GO; GO:0007267; P:cell-cell signaling; TAS.
 CC DR GO; GO:0007139; P:histogenesis and organogenesis; TAS.
 CC DR GO; GO:0008284; P:positive regulation of cell proliferation; TAS.
 CC DR GO; GO:0007165; P:signal transduction; TAS.
 CC DR InterPro; IPR000859; CUB.
 CC DR InterPro; IPR00421; FAS8.C.
 CC DR InterPro; IPR008979; Gal bind like.
 CC DR InterPro; IPR000998; MAM_domain.
 CC DR Pfam; PF00431; CUB; 2.
 CC DR Pfam; PF00754; F5_P8_type_C; 2.
 CC DR Pfam; PF00629; MAM; 1.
 CC DR PRINTS; PRO0020; MAMDOMAIN.
 CC DR SMART; SM00042; CUB; 2.
 CC DR SMART; SM00231; FAS8C; 2.
 CC DR SMART; SM00137; MAM; 1.
 CC DR PROSITE; PS01180; CUB; 2.
 CC DR PROSITE; PS01285; FAS8C_1; 2.
 CC DR PROSITE; PS01286; FAS8C_2; 2.
 CC DR PROSITE; PS00022; FAS8C_3; 2.
 CC DR PROSITE; PS00740; MAM_1; 1.
 CC DR PROSITE; PS00060; MAM_2; 1.
 CC DR Angiogenesis; Transmembrane; Glycoprotein; Neuron; Signal; Repeat;
 CC Receptor; Alternative splicing; 3D-structure.
 CC KM
 CC FT CHAIN 1 21 NEURAPILIN-1.
 CC FT DOMAIN 22 923 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 22 856 POTENTIAL.
 CC FT DOMAIN 857 879 CYTOPLASMIC (POTENTIAL).
 CC FT DOMAIN 880 923 CUB 1.
 CC FT DOMAIN 27 141 CUB 2.
 CC FT DOMAIN 147 265 F5/8 TYPE C 1.
 CC FT DOMAIN 275 424 F5/8 TYPE C 2.
 CC FT DOMAIN 431 583 MAM.
 CC FT DOMAIN 645 811 PROBABLE.
 CC FT DISULFID 27 54 PROBABLE.
 CC FT DISULFID 82 104 PROBABLE.
 CC FT DISULFID 147 173 PROBABLE.
 CC FT DISULFID 206 228 PROBABLE.
 CC FT DISULFID 275 424 BY SIMILARITY.
 CC FT DISULFID 431 583 BY SIMILARITY.
 CC FT CARBOHYD 150 583 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 842 842 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VASAPLIC 642 644 EEP -> GIK (in isoform 2).
 FT VASAPLIC 645 923 /FtId=VSP_004339.
 FT VASAPLIC 645 923 Missing (in isoform 2).
 FT CONFLICT 26 26 /FtId=VSP_004340.
 FT CONFLICT 749 749 K -> E (IN REF. 1).
 FT CONFLICT 855 855 D -> H (IN REF. 2).
 FT CONFLICT 855 855 E -> D (IN REF. 2).
 SQ SEQUENCE 923 AA; 103120 MW; ADKADCA484955D57 CRC64;
 Query Match 25.6%; Score 221; DB 1; Length 923;
 Best local similarity 41.2%; Pred. No. 6.3e-15;
 Matches 47; Conservative 18; Mismatches 47; Indels 2; Gaps 2;
 QY 40 QCGTWKAEIGITSPVPSKYPDPDECIYIEAA-PROCTIEYFDEKSIPEWCKE 98
 DB 26 KCGDTIKESGYLISGYPHSHYSEKCEWLQAPDPYQIMNFNHPLEDR-DCKY 84
 QY 99 DHIYRQGFSPSPILIGRPGCGQNPVYKSSGRFLWIKFPADGEISMGFSARY 152
 DB 85 DYVEVFDGENENGHFRGKFCGKIAPPVVS8GPFLEFIKIVSDYTHGAGFSIRY 138
 RESULT 8
 ID BMP1 HUMAN STANDARD; PRT; 986 AA.
 AC P13497; Q13292; Q13872; Q14874; Q99421; Q99423; Q9UL38;
 DT 01-JAN-1990 (Ref. 13, Created)
 DT 16-OCT-2001 (Ref. 40, Last sequence update)
 DT 15-MAR-2004 (Ref. 43, Last annotation update)
 DE Bone morphogenetic protein 1 precursor (EC 3.4.24.19) (BMP-1)
 DE (Procollagen C-proteinase) (PCP) (Mammalian tollloid protein) (MTld).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP TISSUE=FROM N.A. (ISOFORM BMP1-3).
 RC TISSUE=Skin;
 RX MEDLINE=96209868; PubMed=664539;
 RA Li S.W., Sieron A.L., Ferrata A., Hojima Y., Arnold W.V.,
 RA Prockop D.J.;
 RT "The C-proteinase that processes procollagens to fibrillar collagens
 RT is identical to the protein previously identified as bone morphogenic
 RT protein-1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:5127-5130(1996).
 RP SEQUENCE FROM N.A. (ISOFORM BMP1-1).
 RX MEDLINE=89072730; PubMed=3201241;
 RA Wozney J.M., Rosen V., Celeste A.J., Mittleman L.M., Whitters M.J.,
 RA Kriz R.W., Hewick R.M., Wang E.A.;
 RT "Novel regulators of bone formation: molecular clones and
 RT activities";
 RL Science 242:1528-1534(1988).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS BMP1-4; BMP1-5 AND BMP1-6).
 RC TISSUE=Placenta;
 RX MEDLINE=98160316; PubMed=9500680;
 RA Janitz M., Heiser V., Boettcher U., Landt O., Lauster R.;
 RT "Three alternatively spliced variants of the gene coding for the human
 RT bone morphogenetic protein-1.";
 RL J. Mol. Med. 76:141-146(1998).
 RN [4]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORMS BMP1-3 AND BMP1-7).
 RC TISSUE=Placenta;
 RX MEDLINE=95096114; PubMed=7796260;
 RA Takahara K., Lyons G.E., Greenspan D.S.;
 RT "Bone morphogenetic protein-1 and a mammalian tollloid homologue (mtld)
 RT are encoded by alternatively spliced transcripts which are
 RT differentially expressed in some tissues.";
 RL J. Biol. Chem. 269:32572-32578(1994).
 RN [5]

Db 478 DSGQIQSNPYDDYDRPSKCEWRIIVSGYVGLSF-QVFEIERHDSQAVDYLEVRDGLS 536
 Qy 109 GFSPIIFRCGQONPPVITKSGRFLMTFRADGELESNGFSARY 152
 Db 537 ENSPIIFRCGQKEDIRSTSNLMKRVSDGTVMKAGFANF 580

RESULT 10
 TLD_DROME STANDARD; PRT; 1057 AA.

AC P25723; Q9VC46;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Dorsal-ventral patterning tolloid protein precursor (BC 3.4.24.-).
 GN TLD OR CG6868.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCBI_Taxid=7227;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Canlon-S;
 RX MEDLINE=92034970; PubMed=1840509;
 RA Shmell M.J., Ferguson E.L., Childs S.R., O'Connor M.B.;
 RT "The Drosophila dorsal-ventral patterning gene tolloid is related to
 human bone morphogenetic protein 1.";
 RL Cell 67:469-481 (1991).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC MEDLINE=95324373; PubMed=7600963;
 RA Finelli A.L., Bossie C.A., Xie T., Padgett R.W.;
 RT "Mutational analysis of the Drosophila tolloid gene, a human BMP-1
 homolog.";
 RL Development 120:861-870 (1994).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-U., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu I., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolhakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A.L., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Eranclisela C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland J., Wei M.-H., Idegawa C.,
 RA Jatalai M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laske P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskeirn D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter G., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195 (2000).
 CC -1- FUNCTION: Required for normal dorsal development. TLD may interact
 CC physically with Dp-C protein.
 CC -1- MISCELLANEOUS: Mutations in TLD lead to a partial transformation
 CC of dorsal ectoderm into ventral ectoderm.
 CC -1- SIMILARITY: Belongs to peptidase family M12A.
 CC -1- SIMILARITY: Contains 2 EGF-like domains.
 CC -1- SIMILARITY: Contains 5 CUB domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M76976; AAA8491.1; -.
 DR EMBL: U04239; AAC6482.1; -.
 DR EMBL: AB03749; AAF56323.1; -.
 DR PIR: A39288; A39288.
 DR HSRP: P00742; 1HCG.
 DR MEROPS: M12.010; -.
 DR Pfam: PF00431; CUB; 5.
 DR Pfam: PF00008; EGF; 2.
 DR PRINTS: PR00480; ASTACIN.
 DR SMART: SM00442; CUB; 5.
 DR SMART: SM00179; EGF_CA; 2.
 DR SMART: SM00235; ZDMC; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 2.
 DR PROSITE: PS01180; CUB_5; FALSE_NEG.
 DR PROSITE: PS00022; EGF_1; FALSE_NEG.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS00026; EGF_3; 2.
 DR PROSITE: PS01187; EGF_CA; 2.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 KW Developmental protein; Hydrolyase; Metalloprotease; Zinc; Glycoprotein;
 KW EGF-like domain; Calcium; Signal; Repeat; Zymogen.
 FT SIGNAL 1 27
 FT PROPEP 28 126
 FT CHAIN 127 1057
 FT
 FT DOMAIN 127 329
 FT DOMAIN 330 467
 FT DOMAIN 468 580
 FT DOMAIN 581 621
 FT DOMAIN 624 742
 FT DOMAIN 743 783
 FT DOMAIN 787 899
 FT DOMAIN 900 1016
 FT DOMAIN 1016 1221
 FT METAL 1222 222
 FT ACT SITE 225 225
 FT METAL 231 231
 FT METAL 235 237
 FT SITE 315 317
 FT SITE 317


```

FT DISULFID 330 380 BY SIMILARITY.
FT DISULFID 407 429 BY SIMILARITY.
FT DISULFID 468 495 BY SIMILARITY.
FT DISULFID 522 544 BY SIMILARITY.
FT DISULFID 585 596 BY SIMILARITY.
FT DISULFID 592 605 BY SIMILARITY.
FT DISULFID 607 620 BY SIMILARITY.
FT DISULFID 624 652 BY SIMILARITY.
FT DISULFID 683 706 BY SIMILARITY.
FT DISULFID 747 758 BY SIMILARITY.
FT DISULFID 754 767 BY SIMILARITY.
FT DISULFID 787 813 BY SIMILARITY.
FT DISULFID 840 862 BY SIMILARITY.
FT DISULFID 900 930 BY SIMILARITY.
FT DISULFID 957 979 BY SIMILARITY.
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 431 431 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 533 533 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 634 634 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 667 667 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 781 781 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 854 854 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 908 908 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1057 AA; 120575 MW; 76F4BSAEB7996FBA CRC64;

```

```

Query Match 25.2%; Score 217.5; DB 1; Length 1057;
Best Local Similarity 38.4%; Pred. No. 1.7e-14;
Matches 43; Conservative 21; Mismatches 47; Indels 1; Gaps 1;

```

```

Qy 39 VQCGTWTHAGGIFTSPNYSKYPPDECIYITAEARQCIIEYFDEKYSIESWEKPF 98
Db 466 VVCGGDLKLTQDSIDSNYMDYKPECVWRITAPDNHVALKE-QSFELEKRGDCAY 524
Qy 99 DHIEFROGPFSPFIIGFCGQNPVYIKSSGRFLWIKFPADEGLSWSGFSFA 150
Db 525 DFVHRIDGNHSDSRIGFCGDKLPNIKTRSNQMIRFVSDSVQKLGFSR 576

```

```

RESULT 11
NRPI RAT STANDARD; PRT; 922 AA.
AC 09QWJ3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neutrophilin-1 precursor (Vascular endothelial cell growth factor 165
receptor).
GN NRPI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=97433085; PubMed=9288754;
RA Kolodkin A.L., Levengood D.V., Rowe E.G., Tai Y.-T., Giger R.D.,
RA Ginty D.D.;
RT "Neutrophilin is a semaphorin III receptor.";
RL Cell 90:753-762(1997).
CC -1- FUNCTION: Receptor involved in the development of the
cardiovascular system, in angiogenesis, in the formation of
certain neuronal circuits and in organogenesis outside the nervous
system. It mediates the chemorepulsive activity of semaphorins. It
binds to semaphorin 3A, the PLGF-2 isoform of PGP, the VEGF-165
isoform of VEGF and VEGF-8. Coexpression with KDR results in
increased VEGF-165 binding to KDR as well as increased chemotaxis.
CC It may regulate VEGF-induced angiogenesis (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Found in the embryonic nervous system.
CC -1- SIMILARITY: Belongs to the neutrophilin family.
CC -1- SIMILARITY: Contains 2 CUB domains.

```

```

CC -1- SIMILARITY: Contains 2 F5/8 type C domains.
CC -1- SIMILARITY: Contains 1 MAM domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
-----

```

```

DR EMBL; AF016296; AAC3337.1; -.
DR HSSP; P12259; IC2T.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR000421; FA58 C.
DR InterPro; IPR008979; Gal Bind like.
DR InterPro; IPR009598; MAM_domain.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR Pfam; PF00629; MAM; 1.
DR PRINTS; PR00020; MAMDOMAIN.
DR SMART; SM00042; MAMDOMAIN.
DR SMART; SM00231; FA58C; 2.
DR SMART; SM0137; MAM; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS00022; FA58C_3; 2.
DR PROSITE; PS00740; MAM_1; 1.
DR PROSITE; PS00600; MAM_2; 1.
DR PROSITE; PS00600; MAM_2; 1.
DR Angiogenesis; Transmembrane; Glycoprotein; Neutrone; Signal; Repeat;
KW Receptor.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 922 NEUTROPHILIN-1.
FT DOMAIN 22 855 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 856 880 POTENTIAL.
FT DOMAIN 881 922 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 27 141 CUB 1.
FT DOMAIN 147 265 CUB 2.
FT DOMAIN 275 424 F5/8 TYPE C 1.
FT DOMAIN 431 583 F5/8 TYPE C 2.
FT DOMAIN 645 811 MAM.
FT DISULFID 27 54 PROBABLE.
FT DISULFID 82 104 PROBABLE.
FT DISULFID 147 173 PROBABLE.
FT DISULFID 206 228 PROBABLE.
FT DISULFID 275 424 BY SIMILARITY.
FT DISULFID 431 583 BY SIMILARITY.
FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 841 841 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 922 AA; 103082 MW; CC6F82AD098B0F2E CRC64;

```

```

Query Match 25.2%; Score 217; DB 1; Length 922;
Best Local Similarity 40.2%; Pred. No. 1.6e-14;
Matches 47; Conservative 20; Mismatches 48; Indels 2; Gaps 2;

```

```

Qy 37 KSVQCGTWTHAGGIFTSPNYSKYPPDECIYITAEARQCIIEYFDEKYSIESWEK 95
Db 23 RSPKCGGTYIKIENPGYITSPGYSHYSPSEKCEWLQAPBPYQRIIMNFPHDLEDR-D 81
Qy 96 CKFDHIEVRNDGPPFPSPFIIGFCGQNPVYIKSSGRFLWIKFPADEGLSWSGFSAR 152
Db 82 CKDYVEVLDGNEGGRLWKGKFGKLPSPVSSGPFIFKVFSDYETHGAGFSIRY 138

```

```

RESULT 12
NRPI MOUSE STANDARD; PRT; 923 AA.
AC P97333;
DT 01-NOV-1997 (Rel. 35, Created)

```

DT 01-NOV-1997 (Rel. 35, last sequence update)
 DT 28-FEB-2003 (Rel. 41, last annotation update)
 DE Neuropilin-1 precursor (A5 protein).
 GN NRP1 OR NRP.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Embryonic Brain;
 RX MEDLINE=96353149; PubMed=8748368;
 RA Kanakami A., Kiteukawa T., Takagi S., Fujisawa H.;
 RT "Developmentally regulated expression of a cell surface protein,
 RT neuropilin, in the mouse nervous system.";
 RL J. Neurobiol. 29:1-17(1996).
 CC -1- FUNCTION: Receptor involved in the development of the
 CC cardiovascular system, in angiogenesis, in the formation of
 CC certain neuronal circuits and in organogenesis outside the nervous
 CC system. It mediates the chemorepellant activity of semaphorins. It
 CC binds to semaphorin 3A, the p162-2 isoform of RGF, the VEGF-165
 CC isoform of VEGF and VEGF-B. Coexpression with KDR results in
 CC increased VEGF-165 binding to KDR as well as increased chemotaxis.
 CC It may regulate VEGF-induced angiogenesis (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Nervous system.
 CC -1- SIMILARITY: Belongs to the neuropilin family.
 CC -1- SIMILARITY: Contains 2 CUB domains.
 CC -1- SIMILARITY: Contains 2 F5/8 type C domains.
 CC -1- SIMILARITY: Contains 1 MAM domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D50086; BAA08789.1; -;
 DR HSSP; P12259; 1CZT.
 DR MGD; MGI:106206; Nrp.
 DR GO; GO:0017154; F:semaphorin receptor activity; IGI.
 DR InterPro; IPR000859; CUB.
 DR InterPro; IPR000421; FAS8_C.
 DR InterPro; IPR008979; Gal_Bind_like.
 DR InterPro; IPR000998; MAM_domain.
 DR Pfam; PF00431; CUB; 2.
 DR Pfam; PF00754; F5_F8_type_C; 2.
 DR Pfam; PF00629; MAM; 1.
 DR PRINTS; PR00020; MAMDOMAIN.
 DR SMART; SM00042; CUB; 2.
 DR SMART; SM00231; FAS8C; 2.
 DR SMART; SM00137; MAM; 1.
 DR PROSITE; PS01180; CUB; 2.
 DR PROSITE; PS01285; FAS8C_1; 2.
 DR PROSITE; PS01286; FAS8C_2; 2.
 DR PROSITE; PS50022; FAS8C_3; 2.
 DR PROSITE; PS00740; MAM_1; 1.
 DR PROSITE; PS50060; MAM_2; 1.
 DR Angiogenesis; Transmembrane; Glycoprotein; Neurone; Signal; Repeat;
 KW Receptor.
 KM
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 923 NEUROPIILIN-1.
 FT DOMAIN 22 856 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 857 879 POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 880 923
 FT DOMAIN 27 141 CUB 1.
 FT DOMAIN 147 265 CUB 2.
 FT DOMAIN 275 424 F5/8 TYPE C 1.
 FT DOMAIN 431 583 F5/8 TYPE C 2.
 FT DOMAIN 645 811 MAM.
 FT DISULFID 27 54 PROBABLE.

FT DISULFID 82 104 PROBABLE.
 FT DISULFID 147 173 PROBABLE.
 FT DISULFID 206 228 PROBABLE.
 FT DISULFID 275 424 BY SIMILARITY.
 FT DISULFID 431 583 BY SIMILARITY.
 FT DISULFID 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 842 842 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 923 AA; 103020 MW; 06448BA170796808 CRC64;
 Query Match 25.1%; Score 216; DB 1; Length 923;
 Best Local Similarity 40.2%; Pred. No. 2.1e-14;
 Matches 47; Conservative 20; Mismatches 48; Indels 2; Gaps 2;
 QY 37 KSVGCGTMTAEGGIFTPSPVPSKYPDRREGIYIEA-APRQCIELYDEKYSIEPSWE 95
 DB 23 RSDKGGTGIKENPEYLTSPGPHSHSEKCEWLIQAPYQRIILINPNPHFDEDR-D 81
 QY 96 CKFDHIEVRDGPFGSPPIIGFCGQONPVIKSGSRPLMKFPADGELSMGFSARY 152
 DB 82 CKYDVEYIDENBEGRLMGKRCRIAPSVSSGPFLPKFVSDVETHAGFSIRY 138
 RESULT 13
 BMDP STRPU
 ID BMDP STRPU STANDARD; PRT; 639 AA.
 AC P98069;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, last annotation update)
 DE Bone morphogenetic protein 1 homolog precursor (EC 3.4.24.-) (SUBMP).
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidea; Echinoidea; Echinacea; Echinoida; Strongylocentrotidae;
 OC Strongylocentrotus.
 OC NCBI_TaxID=7668;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=94215496; PubMed=8162855;
 RA Hwang S.P., Partin U.S., Lemariz W.J.;
 RT "Characterization of a homolog of human bone morphogenetic protein 1
 RT in the embryo of the sea urchin, Strongylocentrotus purpuratus.";
 RL Development 120:559-568(1994).
 CC -1- TISSUE SPECIFICITY: Ectodermal and primary mesenchym cells in
 CC hatched blastula.
 CC -1- DEVELOPMENTAL STAGE: Embryo; highest level before spiculogenesis.
 CC -1- SIMILARITY: Belongs to peptidase family M12A.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 2 CUB domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L23838; AAA30081.1; -;
 DR HSSP; P00736; LAPQ.
 DR MEROPS; M12.005; -;
 DR InterPro; IPR000152; Asx_hydroxyl_S.
 DR InterPro; IPR000859; CUB.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR006025; Pept_M_zn_BS.
 DR InterPro; IPR006026; Peptidase_M.
 DR InterPro; IPR001506; Peptidase_M12A.
 DR Pfam; PF01400; Ascatin_1.
 DR Pfam; PF00431; CUB; 2.

DR Pfam; PF00008; EGF; 1.
 DR PRINTS; PRO0480; ASTACIN.
 DR SMART; SM00042; CUB; 2.
 DR SMART; SM00179; EGF_CA; 1.
 DR SMART; SM00235; ZmC; 1.
 DR PROSITE; PS00010; ASX HYDROXYL; 1.
 DR PROSITE; PS01180; CUB; 2.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00026; EGF_3; 1.
 DR PROSITE; PS01187; EGF_CA; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR Developmental protein; Repeat; Hydrolase; Protease; Zinc; Embryo;
 KM Metalloprotease; EGF-like domain; Calcium; Signal.
 FT SIGNAL 1 23
 FT PROPEP 24 109
 FT CHAIN 110 639
 FT DOMAIN 110 306
 FT DOMAIN 307 419
 FT DOMAIN 420 531
 FT DOMAIN 532 573
 FT METAL 197 197
 FT ACT_SITE 198 198
 FT METAL 201 201
 FT METAL 207 207
 FT DISULFID 536 548
 FT DISULFID 544 557
 FT DISULFID 559 572
 FT CARBOHYD 122 122
 FT CARBOHYD 140 140
 FT CARBOHYD 317 317
 FT CARBOHYD 455 455
 SQ SEQUENCE 639 AA; 71893 MW; 59307B6587894AD CRC64;
 Query Match 24.9%; Score 215; DB 1; Length 639;
 Best Local Similarity 39.3%; Pred. No. 1.7e-14;
 Matches 44; Conservative 19; Mismatches 47; Indels 2; Gaps 2;
 QY 41 CGTWTAKHAGGIFTPNPNSKYPPEPCREYITTEAARPCICELFYDEKYSIEPSWECKPDH 100
 DB 420 CGGHLER-ESGTLQSPNYPDDHPHSEKCVLITMPANTYVGLSF-QSFEIEHETCIYDV 477
 QY 101 IEVRDGPGEFSPITIGRFGQGNQNPVIRKSGRFLTKFADGELMGSARV 152
 DB 478 VEVADGHEDTSPHLIRYICGYFTLPDICKSTGNKMWTFVDSGVNKGFSADP 529
 RESULT 14
 PCOL_HUMAN STANDARD; PRT; 449 AA.
 AC QLSL13; Q14550; (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Procollagen C-proteinase enhancer protein precursor (PCPE) (Type I
 DE procollagen COOH-terminal proteinase enhancer) (Type I procollagen C-
 DE proteinase enhancer protein).
 GN PCOLCE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=95014462; PubMed=7523404;
 RA Takahara K., Kessler E., Biniaminov L., Brusel M., Eddy R.L.,
 RA Jani-Sait S., Shows T.B., Greenspan D.S.;
 RT "Type I procollagen COOH-terminal proteinase enhancer protein:
 RT identification, primary structure, and chromosome localization of the
 RT cognate human gene (PCOLCE).";
 RL J. Biol. Chem. 269:26280-26285 (1994).
 RN [2]

RP REVISIONS TO 56; 154 AND 373.
 RA Kessler E.;
 RL Unpublished observations (FEB-2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RA Hirahara I., Syoutfuda K., Harada K., Tomita M., Urakami K., Terai H.,
 RA Morisaki N., Saito Y.;
 RT "Smooth muscle cell derived procollagen C-protease enhancer protein.";
 RL Cell Struct. Funct. 21:662-662(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99018118; PubMed=9799793;
 RA Gloeckner G., Scherer S., Schattevoy R., Boright A.P., Weber J.,
 RA Tsui L.-C., Rosenthal A.;
 RT "Large-scale sequencing of two regions in human chromosome 7q22:
 RT analysis of 650 kb of genomic sequence around the EPO and CUTL1 loci
 RT reveals 17 genes.";
 RL Genome Res. 8:1060-1073(1998).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=99134301; PubMed=9933570;
 RA Scott I.C., Clark T.G., Takahara K., Hoffman G.G., Greenspan D.S.;
 RT "Structural organization and expression patterns of the human and
 RT mouse genes for the type I procollagen COOH-terminal proteinase
 RT enhancer protein.";
 RL Genomics 55:229-234(1999).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey U., Heltan E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7]
 RP MEDLINE SEQUENCE AND CHARACTERIZATION OF INHIBITORY ACTIVITY.
 RX MEDLINE=20092917; PubMed=10625683;
 RA Mot U.D., Thomas C.L., Rosenbach M.T., Takahara K., Greenspan D.S.,
 RA Banda M.J.;
 RT "Post-translational proteolytic processing of procollagen C-terminal
 RT proteinase enhancer releases a metalloproteinase inhibitor.";
 RL J. Biol. Chem. 275:1384-1390(2000).
 CC -1- FUNCTION: Binds to the COOH-terminal propeptide of type I
 CC procollagen and enhances procollagen C-proteinase activity.
 CC -1- FUNCTION: C-terminal processed part of PCPE (CT-PCPE) may have an
 CC metalloproteinase inhibitory activity.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- PMW: C-terminally processed at multiple positions.
 CC -1- SIMILARITY: Contains 2 CUB domains.
 CC -1- SIMILARITY: Contains 1 NTR domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

This Page Blank (usptol)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 18:21:16 ; Search time 21 Seconds
(Without alignments)
723.727 Million cell updates/sec

Title: US-09-887-593A-2
Perfect score: 862
Sequence: 1 MHGRSVLHVSLIILHL.....ADGELESMGFSARVNFPGK 158

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR 78:*

- 1: pirl:*
- 2: pirl:*
- 3: pirl:*
- 4: pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	818	94.9	533	JC7985
2	260	30.2	321	T33161
3	227.5	26.4	3623	T08618
4	227.5	26.4	3623	T09456
5	227	26.3	707	JC2218
6	224	26.0	991	I49540
7	220.5	25.6	986	B58788
8	217.5	25.2	1057	A35288
9	216.5	25.1	1464	S58984
10	211	24.5	730	BMH1
11	211	24.5	823	A58788
12	199.5	23.1	1070	T31069
13	199	23.1	579	UC7629
14	198.5	23.0	927	UC0948
15	197.5	22.9	449	A55362
16	197.5	22.4	402	JH0403
17	187.5	21.8	1594	T30549
18	178.5	20.7	699	I54763
19	172.5	20.0	2403	A53886
20	171.5	19.9	1524	T30337
21	170.5	19.8	597	T31352
22	164.5	19.1	2083	T42721
23	162.5	18.9	504	S56745
24	161.5	18.7	3871	T22812
25	152.5	17.7	705	CHHUR8
26	151	17.5	275	UC6506
27	151	17.5	277	A41735
28	150	17.4	310	S41055
29	149	17.3	1290	A57190

30	148	17.2	276	2	A47290	TS6-6 homolog B94
31	146	16.9	412	2	S72579	hypothetical prote
32	145	16.8	421	2	T25383	hypothetical prote
33	143.5	16.6	288	2	T33324	hypothetical prote
34	143.5	16.6	417	2	T20199	hypothetical prote
35	142.5	16.5	319	2	I51569	IWS.2 protein - Af
36	142.5	16.5	767	2	JC7018	hypothetical prote
37	141.5	16.4	855	2	T30338	membrane-bound arg
38	137	15.9	1004	2	T30348	oviductin (BC 3.4.
39	131	15.2	416	2	T20448	hypothetical prote
40	128	14.8	383	2	T21946	hypothetical prote
41	125	14.5	770	2	T00203	IDL receptor relat
42	123	14.3	770	2	T00204	IDL receptor relat
43	122.5	14.2	317	2	JC7592	spinal cord-derive
44	121.5	14.1	417	2	T29864	hypothetical prote
45	118	13.7	370	2	JC7591	spinal cord-derive

ALIGNMENTS

RESULT 1

JC7985

brain-specific CUB and LDla domain-containing transmembrane protein, BTCL1 protein - M

C/Species: Mus musculus (house mouse)

C/Date: 10-Nov-2003 #sequence_revision 10-Nov-2003 #text_change 24-Nov-2003

C/Accession: JC7985; PC7228

R/Michishita, M.; Ikeda, T.; Nakashiba, T.; Ogawa, M.; Tashiro, K.; Honjo, T.; Doi, K.

Biochem. Biophys. Res. Commun. 306, 680-686, 2003

A/Title: A novel gene, Btcl1, encoding CUB and LDla domains is expressed in restricted

A/Reference number: JC7985; PMID:12810072

A/Accession: JC7985

A/Molecule type: mRNA

A/Residues: 1-533 <MIC>

A/Cross-references: GB:AY138990

A/Experimental source: Cerebellum, 3-week-old

A/Accession: PC7228

A/Molecule type: protein

A/Residues: 1-61 <MT2>

C/Comment: This protein belongs to a new class of brain-specific type I transmembrane p

rtins, and plays roles in formation and/or maintenance of neuronal circuits in the cereb

C/Genetics:

A/Keywords: BTCL1

A/Keywords: BTCL1 protein; CUB domain; LDla domain; transmembrane protein

Query Match 94.9%; Score 818; DB 2; Length 533;

Best Local Similarity 94.2%; Pred. No. 6,1e-76;

Matches 147; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 MHGRSVLHVSLIILHLGATKTKTQTSFQSVQCGTWTKEAGGIFSPNYS 60

DB 1 MYGRSLPHILIASLIILHSSGATKTKTQTSFQSVQCGTWTKEAGGIFSPNYS 60

QY 61 KYPPRECIYIIIEAIPROCIIEYDEKYSIEPSWECKFDHIEVRDGPFGFSPIIRFCQ 120

DB 61 KYPPRECIYIIIEAIPROCIIEYDEKYSIEPSWECKFDHIEVRDGPFGFSPIIRFCQ 120

QY 121 QNPPIYKSGRFLWIKFADGELSMGFSARVNF 156

DB 121 QNPPIYKSGRFLWIKFADGELSMGFSARVNF 156

RESULT 2

T33161

hypothetical protein K035.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C/Accession: T33161

R/Becker, M.; Graves, T.; Fronick, B.

submitted to the EMBL Data Library, May 1998

A/Description: The sequence of C. elegans coamid K035.

A/Reference number: Z1294

A/Accession: T33161

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-321 <BEC>
 A:Cross-references: EMBL:AF067619; PIDN:AACT17565.1; GSPDB:GN00019; CESP:K03E5.1
 A:Experimental source: strain Bristol N2; clone K03E5
 A:Gene: CESP:K03E5.1
 A:Map position: 1
 A:Introns: 25/1; 47/1; 84/2; 112/3; 131/3; 162/1; 197/1; 275/2

Query Match 30.2%; Score 260; DB 2; Length 321;
 Best Local Similarity 43.5%; Pred. No. 6.3e-19;
 Matches 50; Conservative 20; Mismatches 37; Indels 8; Gaps 2;

Db 75 IFTSPNVPKYPDPRECIVITIAAPROCIETYDEKYSIEPSM-----ECKFDHIEVR 104
 135 DGRYGFSPILARFCGDMPRREIRAVSGFLMIRSDSMLEYGCSAEYALVPSK 189

RESULT 3
 T08618
 Intrinsic factor-B12 receptor CUBILIN precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 02-Aug-2002
 C:Accession: T08618
 R:Moestrup, S.K.; Kozzyraki, R.; Kristiansen, M.; Kayser, J.H.; Rasmussen, H.H.; Brault, J. Biol. Chem. 273, 5235-5242, 1998
 A:Title: The intrinsic factor-vitamin B12 receptor and target of teratogenic antibodies
 A:Reference number: Z16459; MUID:98148073; PMID:9478979
 A:Accession: T08618
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-3623 <MOE>
 A:Cross-references: EMBL:AF022247; NID:93834379; PIDN:AACT1661.1; PID:93834380
 A:Gene: CUBILIN
 C:Keywords: intrinsic factor-B12 receptor cubilin; EGF homology
 C:Keywords: egg yolk; endocytosis; glycoprotein; intestine; kidney; peripheral membrane
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-3623/Product: intrinsic factor-B12 receptor CUBILIN #status predicted <MAT>
 F:133-164/Domain: EGF homology <EGF1>
 F:436-467/Domain: EGF homology <EGF>

Query Match 26.4%; Score 227.5; DB 2; Length 3623;
 Best Local Similarity 37.5%; Pred. No. 2.1e-14;
 Matches 45; Conservative 26; Mismatches 46; Indels 3; Gaps 2;

Db 37 KSVQCG--TWTKHAE--GIFTSPNVPKYPDPRECIVITIAAPROCIETYDEKYSIEPS 93
 2213 KSIACGGNVYIHDAISDGYLSPNVPANYPQAEICIMLEAPGGRSIOLOQFEDQNIEDT 2272

Query 94 WECKFDHIEVRDGPFGFSPILGRFCGQONPPVTKSGRFLMKFPADGELBSMGFSARYN 153
 2273 PNCSTNVLELRDGVSDAPILSKFCGTSLPSSQWSSGEVWYLRFRSDNSPTHVGFKAKYS 2332

RESULT 4
 T09456
 Intrinsic factor-B12 receptor Cubilin precursor - human
 C:Species: Homo sapiens (man)
 C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 02-Aug-2002
 C:Accession: T09456
 R:Kozzyraki, R.; Kristiansen, M.; Sliacharoglu, A.; Hansen, C.; Jacobsen, C.; Tommerup, N
 Blood 91, 3593-3600, 1998
 A:Title: The human intrinsic factor-vitamin B12 receptor, cubilin: Molecular characterization.
 A:Reference number: Z16677; MUID:98241400; PMID:9572993
 A:Accession: T09456
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
 A:Residues: 1-3623 <KOZ>
 A:Cross-references: EMBL:AF034611; NID:93929528; PIDN:AA082612.1; PID:93929529
 A:Gene: CESP:K03E5.1
 A:Map position: 10p12
 C:Superfamily: intrinsic factor-B12 receptor cubilin; EGF homology
 C:Keywords: receptor; vitamin B12 uptake
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-3623/Product: intrinsic factor-B12 receptor #status predicted <MAT>
 F:436-467/Domain: EGF homology <EGF>

Query Match 26.4%; Score 227.5; DB 2; Length 3623;
 Best Local Similarity 35.1%; Pred. No. 2.1e-14;
 Matches 46; Conservative 29; Mismatches 47; Indels 9; Gaps 3;

Db 37 KSVQCG--TWTKHAE--GIFTSPNVPKYPDPRECIVITIAAPROCIETYDEKYSIEPS 93
 2213 KSIACGGNVYIHDAISDGYLSPNVPANYPQAEICIMLEAPGGRSIOLOQFEDQNIEDT 2272

Query 94 WECKFDHIEVRDGPFGFSPILGRFCGQONPPVTKSGRFLMKFPADGELBSMGFSARYN 153
 2273 PNCSTNVLELRDGVSDAPILSKFCGTSLPSSQWSSGEVWYLRFRSDNSPTHVGFKAKYS 2332

Db 154 F-----TPGK 158
 2333 IAQCGGRVPCQ 2343

RESULT 5
 JC2218
 Procollagen C-endopeptidase (BC 3.4.24.19) - African clawed frog (clone 22AN)
 N:Alternate names: bone morphogenic protein 1
 C:Species: Xenopus laevis (African clawed frog)
 C>Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 20-Sep-1999
 C:Accession: JC2218
 R:Maeno, M.; Xue, Y.; Wood, T.I.; Ong, R.C.; Kung, H.
 Gene 134, 257-261, 1993
 A:Title: Cloning and expression of cDNA encoding Xenopus laevis bone morphogenetic prot.
 A:Reference number: JC2218; MUID:94085787; PMID:8262384
 A:Accession: JC2218
 A:Molecule type: mRNA
 A:Residues: 1-707 <MAB>
 A:Cross-references: GB:L1249; NID:9406540; PIDN:AAA16313.1; PID:9406541
 C:Comment: This protein induces ectopic cartilage formation in vivo.
 C:Superfamily: procollagen C-endopeptidase; astacin homology; C1r/C1s repeat homology;
 C:Keywords: beta-hydroxyasparagine; glycoprotein; hydrolase; metalloproteinase; zinc
 F:93-284/Domain: astacin homology <AST>
 F:285-397/Region: complement 1r/1s-like repeat
 F:285-394/Domain: C1r/C1s repeat homology <C1R1>
 F:398-510/Region: complement 1r/1s-like repeat
 F:398-507/Domain: C1r/C1s repeat homology <C1R2>
 F:514-550/Domain: EGF homology <EGF>
 F:554-666/Region: complement 1r/1s-like repeat
 F:554-663/Domain: C1r/C1s repeat homology <C1R3>
 F:624-105/295,326/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:176-180,186,235/Binding site: zinc (His, His, Tyr) #status predicted
 F:177/Active site: Glu #status predicted
 F:528/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 26.3%; Score 227; DB 2; Length 707;
 Best Local Similarity 40.2%; Pred. No. 3.7e-15;
 Matches 45; Conservative 18; Mismatches 47; Indels 2; Gaps 2;

Db 41 CGTWTKHAE--GIFTSPNVPKYPDPRECIVITIAAPROCIETYDEKYSIEPSWECKFDH 100
 398 CGGEVKKYDSGHI--QSPNVPDIDRPRKACWKLVSBEFGVIGSF--QSPFIEHNDSCAYDY 455

Query 101 IEVRDGPFGFSPILGRFCGQONPPVTKSGRFLMKFPADGELBSMGFSARYN 152
 456 IEIRQSSSETSLVGRFCGYDKRPDDIKSSTNOLMKFVSDGSLNKAGFSLNY 507

RESULT 6

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 18:23:47 ; Search time 43 Seconds

(without alignments)
1022,449 Million cell updates/sec

Title: US-09-887-593a-2

Perfect score: 862
Sequence: 1 MIHRSVLIHVASLIILHLS.....ADGELMSGFSARNVTPGK 158

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1145568 seqs, 278261457 residues

Total number of hits satisfying chosen parameters: 1145568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10C_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	862	100.0	158	9	US-09-887-593-2
2	851	98.7	526	15	US-10-138-588-38
3	851	98.7	533	13	US-10-041-406-2
4	851	98.7	533	14	US-10-162-435-27
5	479	55.6	130	9	US-09-764-853-653
6	479	55.6	141	10	US-09-764-881-116
7	479	55.6	141	15	US-09-764-881-116
8	479	55.6	141	15	US-10-242-747-116
9	470	54.5	352	12	US-10-276-774-2238
10	470	54.5	525	12	US-10-147-423-114
11	470	54.5	525	12	US-10-145-127-114
12	470	54.5	525	12	US-10-160-503-114
13	470	54.5	525	12	US-10-118-114
14	470	54.5	525	12	US-10-143-993-114
15	470	54.5	525	12	US-10-158-787-114

16	470	54.5	525	12	US-10-140-024-114	Sequence 114, App
17	470	54.5	525	12	US-10-140-808-114	Sequence 114, App
18	470	54.5	525	12	US-10-152-405-114	Sequence 114, App
19	470	54.5	525	12	US-10-127-852A-114	Sequence 114, App
20	470	54.5	525	12	US-10-127-900A-114	Sequence 114, App
21	470	54.5	525	12	US-10-128-685A-114	Sequence 114, App
22	470	54.5	525	12	US-10-131-820A-114	Sequence 114, App
23	470	54.5	525	12	US-10-142-886-114	Sequence 114, App
24	470	54.5	525	12	US-10-146-728-114	Sequence 114, App
25	470	54.5	525	12	US-10-146-768-114	Sequence 114, App
26	470	54.5	525	12	US-10-147-499-114	Sequence 114, App
27	470	54.5	525	12	US-10-157-798-114	Sequence 114, App
28	470	54.5	525	14	US-10-028-072-114	Sequence 114, App
29	470	54.5	525	14	US-10-121-049-114	Sequence 114, App
30	470	54.5	525	14	US-10-123-904-114	Sequence 114, App
31	470	54.5	525	14	US-10-140-470-114	Sequence 114, App
32	470	54.5	525	14	US-10-175-746-114	Sequence 114, App
33	470	54.5	525	14	US-10-176-918-114	Sequence 114, App
34	470	54.5	525	14	US-10-176-921-114	Sequence 114, App
35	470	54.5	525	14	US-10-137-865-114	Sequence 114, App
36	470	54.5	525	14	US-10-140-474-114	Sequence 114, App
37	470	54.5	525	14	US-10-142-431-114	Sequence 114, App
38	470	54.5	525	14	US-10-143-114-114	Sequence 114, App
39	470	54.5	525	14	US-10-140-002-114	Sequence 114, App
40	470	54.5	525	14	US-10-142-419-114	Sequence 114, App
41	470	54.5	525	14	US-10-123-262-114	Sequence 114, App
42	470	54.5	525	14	US-10-142-423-114	Sequence 114, App
43	470	54.5	525	14	US-10-213-181-12	Sequence 12, Appl
44	470	54.5	525	14	US-10-121-050-114	Sequence 114, App
45	470	54.5	525	14	US-10-141-755-114	Sequence 114, App

ALIGNMENTS

RESULT 1
US-09-887-593-2
; Sequence 2, Application US/09887593
; Patent No. US20020161212A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel B.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Leong, Kanun
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Saffran, Douglas C.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: BPC-1: A SECRETED BRAIN-SPECIFIC PROTEIN EXPRESSED AND
; TITLE OF INVENTION: SECRETED BY PROSTATE AND BLADDER CANCER CELLS
; FILE REFERENCE: 1703-017 US1
; CURRENT APPLICATION NUMBER: US/09/887,593
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 09/374,135
; PRIOR FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: 60/095,982
; PRIOR FILING DATE: 1998-08-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-887-593-2

Query Match 100.0%; Score 862; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 44e-87;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIHRSVLIHVASLIILHLSGATKGTGKOTTSFQKSVOCGWTAKAEGGIFTSNYPVS 60
Db 1 MIHRSVLIHVASLIILHLSGATKGTGKOTTSFQKSVOCGWTAKAEGGIFTSNYPVS 60
QY 61 KYPRDRCIYIIEAPQCEIYFDEKYSIIEPSGECKFDHIEVVDGPRGSPITIGRPGQ 120

Db 61 KYPDRECIYIIIEAAROCIELYFDEKYSIEPSWECKFDHIEVRDGFPGFSPILIGRCCQ 120
QY 121 QNPVTKSSGRFLMIKFFADGELESMGFSARYNFTPGK 158
Db 121 QNPVTKSSGRFLMIKFFADGELESMGFSARYNFTPGK 158

RESULT 2
US-10-138-588-38
; Sequence 38, Application US/10138588
; Publication No. US20040018594A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook et al.
; TITLE OF INVENTION: NOVEL ANTIBODIES THAT BIND TO ANTIGENIC POLYPEPTIDES, NUCLEIC ACID
; FILE REFERENCE: 21402-347A
; CURRENT APPLICATION NUMBER: US/10/138,588
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 60/288,395
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/308,901
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/313,388
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/324,757
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/288,900
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 203
; SEQ ID NO 38
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-138-588-38

Query Match 98.7%; Score 851; DB 15; Length 526;
Best Local Similarity 100.0%; Pred. No. 3.4e-85;
Matches 156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHGSRVLIHVASLIILHSGATKKGTEKOTTSVQCKGWTWKAEGGIFTSNPNYS 60
Db 1 MHGSRVLIHVASLIILHSGATKKGTEKOTTSVQCKGWTWKAEGGIFTSNPNYS 60

QY 61 KYPDRECIYIIIEAAROCIELYFDEKYSIEPSWECKFDHIEVRDGFPGFSPILIGRCCQ 120
Db 61 KYPDRECIYIIIEAAROCIELYFDEKYSIEPSWECKFDHIEVRDGFPGFSPILIGRCCQ 120

QY 121 QNPVTKSSGRFLMIKFFADGELESMGFSARYNFTP 156
Db 121 QNPVTKSSGRFLMIKFFADGELESMGFSARYNFTP 156

RESULT 3
US-10-041-406-2
; Sequence 2, Application US/10041406
; Publication No. US20020164705A1
; GENERAL INFORMATION:
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: 39362, A NOVEL CUB DOMAIN CONTAINING
; FILE REFERENCE: 10448-130001
; CURRENT APPLICATION NUMBER: US/10/041,406
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: 60/260,286
; PRIOR FILING DATE: 2001-01-08
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-041-406-2

Query Match 98.7%; Score 851; DB 13; Length 533;
Best Local Similarity 100.0%; Pred. No. 3.4e-85;
Matches 156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHGSRVLIHVASLIILHSGATKKGTEKOTTSVQCKGWTWKAEGGIFTSNPNYS 60
Db 1 MHGSRVLIHVASLIILHSGATKKGTEKOTTSVQCKGWTWKAEGGIFTSNPNYS 60

QY 61 KYPDRECIYIIIEAAROCIELYFDEKYSIEPSWECKFDHIEVRDGFPGFSPILIGRCCQ 120
Db 61 KYPDRECIYIIIEAAROCIELYFDEKYSIEPSWECKFDHIEVRDGFPGFSPILIGRCCQ 120

QY 121 QNPVTKSSGRFLMIKFFADGELESMGFSARYNFTP 156
Db 121 QNPVTKSSGRFLMIKFFADGELESMGFSARYNFTP 156

RESULT 4
US-10-162-435-27
; Sequence 27, Application US/10162435
; Publication No. US20030096305A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Kapeller-Liebermann, Rosana
; APPLICANT: Bandaru, Rajasekhar
; APPLICANT: Leiby, Kevin R.
; TITLE OF INVENTION: NOVEL HUMAN MEMBRANE-ASSOCIATED PROTEIN AND
; FILE REFERENCE: 10448-189001
; CURRENT APPLICATION NUMBER: US/10/162,435
; PRIOR FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 09/836,499
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/US01/12420
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: US 60/197,507
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/891,008
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: PCT/US01/19963
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/214,220
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/860,868
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: PCT/US01/16013
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,674
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/886,429
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: PCT/US01/20055
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/213,963
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 10/041,406
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: PCT/US02/00275
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/260,286
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: US 09/934,268
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/41811
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,612
; PRIOR FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 533

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-435-27

Query Match
Best Local Similarity 98.7%; Score 851; DB 14; Length 533;
Matches 156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHGSSVLIIVASLIILHSGATKGTETKOTTSFQKSVQCGTWTGAEGIFTSNPYS 60
DB 1 MHGSSVLIIVASLIILHSGATKGTETKOTTSFQKSVQCGTWTGAEGIFTSNPYS 60
QY 61 KPPDRECIYITIAAPROCIELYPDEKYSIEPSWECKPHIEVRDGPFGFSPITGRFCQ 120
DB 61 KPPDRECIYITIAAPROCIELYPDEKYSIEPSWECKPHIEVRDGPFGFSPITGRFCQ 120
QY 121 QNPVYIKSSGRFLMIKFPADGELSMGFSARYNFTP 156
DB 121 QNPVYIKSSGRFLMIKFPADGELSMGFSARYNFTP 156

RESULT 5
US-09-764-853-653
; Sequence 653, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ206
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 653
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-853-653

Query Match
Best Local Similarity 55.6%; Score 479; DB 9; Length 130;
Matches 80; Conservative 21; Mismatches 20; Indels 0; Gaps 0;

QY 38 SVQCGTWTGAEGIFTSNPYSKYPDPRECIYITIAAPROCIELYPDEKYSIEPSWECK 97
DB 9 ATQCGIWRSTNSGHHFASPNYPDSYPNKECIYITIAAPROCIELYPDEHYIIEPSFECR 68
QY 98 FDHIEVRDGPFGFSPITGRFCQGNPVIKSSGRFLMIKFPADGELSMGFSARYNFTPG 157
DB 69 FDHIEVRDGPFGFSPITGRFCQGNPVIKSSGRFLMIKFPADGELSMGFSARYNFTPG 128
QY 158 K 158
DB 129 K 129

RESULT 6
US-09-764-881-116
; Sequence 116, Application US/09764881
; Publication No. US20030125246A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 116
; LENGTH: 141
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-09-764-881-116

Query Match
Best Local Similarity 55.6%; Score 479; DB 10; Length 141;
Matches 80; Conservative 21; Mismatches 20; Indels 0; Gaps 0;

QY 38 SVQCGTWTGAEGIFTSNPYSKYPDPRECIYITIAAPROCIELYPDEKYSIEPSWECK 97
DB 9 ATQCGIWRSTNSGHHFASPNYPDSYPNKECIYITIAAPROCIELYPDEHYIIEPSFECR 68
QY 98 FDHIEVRDGPFGFSPITGRFCQGNPVIKSSGRFLMIKFPADGELSMGFSARYNFTPG 157
DB 69 FDHIEVRDGPFGFSPITGRFCQGNPVIKSSGRFLMIKFPADGELSMGFSARYNFTPG 128
QY 158 K 158
DB 129 K 129

RESULT 7
US-09-764-881-116
; Sequence 116, Application US/09764881
; Publication No. US20020086821A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 116
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-881-116

Query Match
Best Local Similarity 55.6%; Score 479; DB 12; Length 141;
Matches 80; Conservative 21; Mismatches 20; Indels 0; Gaps 0;

QY 38 SVQCGTWTGAEGIFTSNPYSKYPDPRECIYITIAAPROCIELYPDEKYSIEPSWECK 97
DB 9 ATQCGIWRSTNSGHHFASPNYPDSYPNKECIYITIAAPROCIELYPDEHYIIEPSFECR 68
QY 98 FDHIEVRDGPFGFSPITGRFCQGNPVIKSSGRFLMIKFPADGELSMGFSARYNFTPG 157
DB 69 FDHIEVRDGPFGFSPITGRFCQGNPVIKSSGRFLMIKFPADGELSMGFSARYNFTPG 128
QY 158 K 158
DB 129 K 129

RESULT 8
US-10-242-747-116
; Sequence 116, Application US/10242747
; Publication No. US20040005577A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207C1
; CURRENT APPLICATION NUMBER: US/10/242,747
; CURRENT FILING DATE: 2002-09-13
; Prior application data removed - refer to PALM or file wrapper
; Prior application number: 09/764,881
; Prior application number: 2001-01-17
; Prior application number: 60/179,065
; Prior application number: 2000-01-31
; Prior application number: 60/180,628
; Prior application number: 2000-02-04
; Prior application number: 60/214,886
```

```
;; PRIOR FILING DATE: 2000-06-28
;; PRIOR APPLICATION NUMBER: 60/217,487
;; PRIOR FILING DATE: 2000-07-11
;; PRIOR APPLICATION NUMBER: 60/225,758
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/220,963
;; PRIOR FILING DATE: 2000-07-26
;; PRIOR APPLICATION NUMBER: 60/217,496
;; PRIOR FILING DATE: 2000-07-11
;; PRIOR APPLICATION NUMBER: 60/225,447
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/218,290
;; PRIOR FILING DATE: 2000-07-14
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 192
;; SOFTWARE: Patent Ver. 2.0
;; SEQ ID NO 116
;; LENGTH: 141
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-242-747-116
```

```
Query Match 55.6%; Score 479; DB 15; Length 141;
Best Local Similarity 66.1%; Pred. No. 8.4e-45;
Matches 80; Conservative 21; Mismatches 20; Indels 0; Gaps 0;
```

```
QY 38 SVCCGTAKAEGIFTSPPVPSKYPDRRCIYIIAAPQCIELYFDEKYSIPECK 97
DB 9 ATCCGIVWRNSNGHFASPPDSDYPNKECIYIIAAPQRIELTFDEHYIEPSECR 68
```

```
QY 98 FDHIEVDGPFSPFIIGRCGQGNPPVIRKSGRFLMIKEFADGEISMGFSAVNFPG 157
DB 69 FDHIEVDGPFSPFLIDRYCGVSPPLINSTGRFMWIKKSDDELEGLGFRAYSTPG 128
```

```
QY 158 K 158
DB 129 K 129
```

```
RESULT 9
US-10-276-774-2238
;; Sequence 2238, Application US/10276774
;; Publication No. US20040053245A1
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq, Inc.
;; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
;; FILE REFERENCE: 21272-030
;; CURRENT APPLICATION NUMBER: US/10/276,774
;; PRIOR FILING DATE: 2002-11-18
;; PRIOR APPLICATION NUMBER: 09/560,875
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 09/496,914
;; PRIOR FILING DATE: 2000-02-03
;; NUMBER OF SEQ ID NOS: 2700
;; SOFTWARE: Custom
;; SEQ ID NO 2238
;; LENGTH: 352
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-276-774-2238
```

```
Query Match 54.5%; Score 470; DB 12; Length 352;
Best Local Similarity 54.9%; Pred. No. 2.6e-43;
Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;
```

```
QY 6 SVLHI--VASLIILHLSGATKKGTSEKOTSEKSVCCGTWKAEAGIFTSPPVPSKYP 63
DB 28 SVLKVLLITVAVVGIAVAAKTDGQGNIGIKHPIATCCGIVWRNSNGHFASPPYDSDP 87
QY 64 PDRECIYIIAAPQCIELYFDEKYSIPECKFDHIEVDGPFSPFIIGRCGQGNP 123
DB 88 PNKECIYIIAAPQRIELTFDEHYIEPSECRFDHIEVDGPFSPFLIDRYCGVSP 147
```

```
QY 124 PVTKSGRFLMIKEFADGEISMGFSAVNFPT 156
DB 148 PLIRSTGRFMWIKKSDDELEGLGFRAYSTPG 180
```

```
RESULT 10
US-10-147-493-114
;; Sequence 114, Application US/10147493
;; Publication No. US20040029217A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Beresini, Maureen
;; APPLICANT: DeForge, Laura
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Gao, Wei-Qiang
;; APPLICANT: Gerritsen, Mary E.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Sherwood, Steven
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Watanabe, Colin K
;; APPLICANT: Wood, William
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P330R1G345
;; CURRENT APPLICATION NUMBER: US/10/147,493
;; PRIOR FILING DATE: 2002-05-17
;; Remaining Prior Application data removed - See File Wrapper or Palm
;; NUMBER OF SEQ ID NOS: 550
;; SEQ ID NO 114
;; LENGTH: 525
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-147-493-114
```

```
Query Match 54.5%; Score 470; DB 12; Length 525;
Best Local Similarity 54.9%; Pred. No. 4.4e-43;
Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;
```

```
QY 6 SVLHI--VASLIILHLSGATKKGTSEKOTSEKSVCCGTWKAEAGIFTSPPVPSKYP 63
DB 8 SVLKVLLITVAVVGIAVAAKTDGQGNIGIKHPIATCCGIVWRNSNGHFASPPYDSDP 67
```

```
QY 64 PDRECIYIIAAPQCIELYFDEKYSIPECKFDHIEVDGPFSPFIIGRCGQGNP 123
DB 68 PNKECIYIIAAPQRIELTFDEHYIEPSECRFDHIEVDGPFSPFLIDRYCGVSP 127
```

```
QY 124 PVTKSGRFLMIKEFADGEISMGFSAVNFPT 156
DB 128 PLIRSTGRFMWIKKSDDELEGLGFRAYSTPG 160
```

```
RESULT 11
US-10-145-127-114
;; Sequence 114, Application US/10145127
;; Publication No. US20040033558A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Beresini, Maureen
;; APPLICANT: DeForge, Laura
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Gao, Wei-Qiang
;; APPLICANT: Gerritsen, Mary E.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Gurney, Austin L.
```



```
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C252
CURRENT APPLICATION NUMBER: US/10/145,127
CURRENT FILING DATE: 2002-05-13
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 114
LENGTH: 525
TYPE: PRT
ORGANISM: Homo Sapien
US-10-145-127-114
```

Query Match 54.5%; Score 470; DB 12; Length 525;
Best Local Similarity 54.9%; Pred. No. 4,4e-43;
Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;

```
QY 6 SVLHI--VASLIILHSATKKGTEKQTSETOKSVOCGWTWKAAEGIFTSPNYSKYP 63
DB 8 SVLKVLITLVVVGIAVNAQKTODGQNGIKHLPATQCGIWRTSNGHFASPNYPDSYP 67
QY 64 PDEECYIIIEAARQCIIEYFDEKYSIEPSWECKFDHIEVRDGPFGSPPIIGFCGQNP 123
DB 68 PNEECYIIIEAARQRIELTFDEHYIIEPSFECRFDHLIEVRDGPFGSPFLIDRYCGVXSP 127
QY 124 PVIKSSGRFLMIKFFADGELSMGFSARYNFTP 156
DB 128 PLIRSTGRFMWIKFSSDELEGLGFRAXYSFTP 160
```

RESULT 12
US-10-160-503-114
; Sequence 114, Application US/10160503
; Publication No. US20040033559A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C446
CURRENT APPLICATION NUMBER: US/10/160,503
CURRENT FILING DATE: 2002-05-30
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 114
LENGTH: 525
TYPE: PRT
ORGANISM: Homo Sapien
US-10-160-503-114

Query Match 54.5%; Score 470; DB 12; Length 525;
Best Local Similarity 54.9%; Pred. No. 4,4e-43;

Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;
QY 6 SVLHI--VASLIILHSATKKGTEKQTSETOKSVOCGWTWKAAEGIFTSPNYSKYP 63
DB 8 SVLKVLITLVVVGIAVNAQKTODGQNGIKHLPATQCGIWRTSNGHFASPNYPDSYP 67
QY 64 PDEECYIIIEAARQCIIEYFDEKYSIEPSWECKFDHIEVRDGPFGSPPIIGFCGQNP 123
DB 68 PNEECYIIIEAARQRIELTFDEHYIIEPSFECRFDHLIEVRDGPFGSPFLIDRYCGVXSP 127
QY 124 PVIKSSGRFLMIKFFADGELSMGFSARYNFTP 156
DB 128 PLIRSTGRFMWIKFSSDELEGLGFRAXYSFTP 160

RESULT 13
US-10-143-118-114
; Sequence 114, Application US/10143118
; Publication No. US20040038335A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C228
CURRENT APPLICATION NUMBER: US/10/143,118
CURRENT FILING DATE: 2002-05-09
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 114
LENGTH: 525
TYPE: PRT
ORGANISM: Homo Sapien
US-10-143-118-114

Query Match 54.5%; Score 470; DB 12; Length 525;
Best Local Similarity 54.9%; Pred. No. 4,4e-43;
Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;

```
QY 6 SVLHI--VASLIILHSATKKGTEKQTSETOKSVOCGWTWKAAEGIFTSPNYSKYP 63
DB 8 SVLKVLITLVVVGIAVNAQKTODGQNGIKHLPATQCGIWRTSNGHFASPNYPDSYP 67
QY 64 PDEECYIIIEAARQCIIEYFDEKYSIEPSWECKFDHIEVRDGPFGSPPIIGFCGQNP 123
DB 68 PNEECYIIIEAARQRIELTFDEHYIIEPSFECRFDHLIEVRDGPFGSPFLIDRYCGVXSP 127
QY 124 PVIKSSGRFLMIKFFADGELSMGFSARYNFTP 156
DB 128 PLIRSTGRFMWIKFSSDELEGLGFRAXYSFTP 160
```

RESULT 14
US-10-144-993-114
; Sequence 114, Application US/10144993
; Publication No. US20040038336A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

```
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C261
CURRENT FILING DATE: 2002-05-13
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 114
LENGTH: 525
TYPE: PRT
ORGANISM: Homo Sapien
US-10-144-993-114

Query Match      54.5%; Score 470; DB 12; Length 525;
Best Local Similarity 54.9%; Pred. No. 4.4e-43;
Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;

QY 6 SVLHI--VASLIILHSGATKKGTEKQTSETQSVQCGTWTGHAEGIFTSPNYSKYP 63
DB 8 SVLKVLITLVYVEGIAVAQKTDGQNGIKHLPATQCGIWTSTNGHFASPNYDPSYP 67
QY 64 PDRECIYITFAAPROCIETLYFDEKYSIEPSWECKPHIEVRDGPFGSPPIGFCGQONP 123
DB 68 PNKECIYITFAAPROCIETLYFDEKYSIEPSWECKPHIEVRDGPFGSPPIGFCGQONP 127
QY 124 PVTKSSGRFLWIKFPADGELSMGFSARYNFTP 156
DB 128 PLIRSTGRFWIKFSSDELEGLGFRKYSFIP 160

RESULT 15
US-10-158-787-114
Sequence 114, Application US/10158787
Publication No. US20040039164A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C449
CURRENT FILING DATE: 2003-04-03
```

```
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 114
LENGTH: 525
TYPE: PRT
ORGANISM: Homo Sapien
US-10-158-787-114

Query Match      54.5%; Score 470; DB 12; Length 525;
Best Local Similarity 54.9%; Pred. No. 4.4e-43;
Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;

QY 6 SVLHI--VASLIILHSGATKKGTEKQTSETQSVQCGTWTGHAEGIFTSPNYSKYP 63
DB 8 SVLKVLITLVYVEGIAVAQKTDGQNGIKHLPATQCGIWTSTNGHFASPNYDPSYP 67
QY 64 PDRECIYITFAAPROCIETLYFDEKYSIEPSWECKPHIEVRDGPFGSPPIGFCGQONP 123
DB 68 PNKECIYITFAAPROCIETLYFDEKYSIEPSWECKPHIEVRDGPFGSPPIGFCGQONP 127
QY 124 PVTKSSGRFLWIKFPADGELSMGFSARYNFTP 156
DB 128 PLIRSTGRFWIKFSSDELEGLGFRKYSFIP 160

Search completed: May 18, 2004, 18:29:35
Job time : 44 secs
```